International School for Advanced Studies

condensed matter theory sector

Ph.D. thesis.

Quantum Monte Carlo studies of soft Bosonic systems and Minimum Energy Pathways.

Author:
Sebastiano Saccani

Supervisors:
Saverio Moroni

April 4, 2013

Via Bonomea 265, 34136 Trieste - ITALY

Abstract

In this thesis, we make use of Monte Carlo techniques to address two rather different subjects in condensed matter physics.

The first study deals with the characterization of a relatively novel and elusive phase of matter, the so-called supersolid, in which crystalline order and dissipationless flow coexist. While supersolidity is a well studied phenomenology in lattice models, we will be working here in continuous space, where much fewer results are available. Specifically, we study a soft core Bosonic system, quantum analog of thoroughly studied classical models, which displays an unambiguous supersolid phenomenology. In this system such a behavior is not obtained through Bose Condensation of lattice defects, but rather it is mean field in character. By computer simulations we characterize many properties of the system: of these, the most prominent are the phase diagram of the system and its excitation spectrum. This study is loosely related to the ultracold atom experimental field, as it is speculated that interparticle potential pertaining to the same class of the one employed here may be realized in this context. After the recent (and apparently definitive) ruling out of supersolidity effects in ^{4}He, it seems fair to state that ultracold atoms are the most promising candidate for the observation of this phenomenology. In this section we employ our own implementation of the worm algorithm on the continuum.

The second part of this thesis is instead related to electronic structure, more specifically to the study of minimum energy pathways of reactions calculated via quantum Monte Carlo methods. In particular, we aim at assessing the computational feasibility and the accuracy of determining the most significant geometries of a reaction (initial/final and transition state) and its energy barrier via these stochastic techniques. To this end, we perform calculations on a set of simple reactions and compare the results with density functional theory and high level quantum chemistry calculations. We show that the employed technique indeed performs better than density functional for both geometries and energy barrier. Therefore our methodology is a good candidate to study reactions in which an high accuracy is needed, but it is not possible to employ high level quantum chemistry methods due to computational limitations. We believe that this study is significant also because of its systematic use of forces from Monte Carlo simulations. Although several studies have addressed various aspects of the problem of computing forces within quantum Monte Carlo accurately and efficiently, there is little awareness that such estimators are in fact mature, and consequently there are very few studies which actually employ them. We hope to show

here that these estimators are actually ready to be used and provide good results. In this section we have mainly developed interfaces for existing Quantum Monte Carlo codes.

Contents

Chapter 1

Soft disk Bosons

1.1 Introduction: supersolidity

We deal in this chapter with an hypothetical phase of matter, displaying simultaneously crystalline order and dissipation-less flow, which is a subject of long standing interest for theorist and experimenters [1–5]. Such a phase, named supersolid, should feature unusual properties, such as a nonclassical momentum of inertia , but has so far escaped unambiguous experimental observation. Crystalline order is determined by spontaneous (i.e. not externally induced) breaking of translational symmetry. This is reflected in periodic fluctuations (long-range order) of the density $\rho(\mathbf{r})$, such that $\rho(\mathbf{r}) = \rho(\mathbf{r} + \mathbf{T})$, $\mathbf{T}$ being a lattice vector. The order parameter for the breaking of translational symmetry is the static structure factor $S(\mathbf{k})$, defined through the Fourier transform of the density

$$S(\mathbf{k}) = \frac{|\rho_{\mathbf{k}}|^2}{N}; \quad \rho_{\mathbf{k}} = \int e^{-i\mathbf{k}\cdot\mathbf{r}}\rho(\mathbf{r}), \tag{1.1}$$

which in the solid phase displays peaks in the correspondence of reciprocal lattice vectors $\mathbf{Q}$, such that $\mathbf{Q} \cdot \mathbf{T} = 2\pi n \ \forall \ \mathbf{T}$, n being an integer number. Dissipationless flow, as the name suggests, is related to the ability of a liquid to flow without friction (superfluidity), and to the concept of nonclassical momentum of inertia. It is common practice to treat the superfluid phenomenology in a two liquid model: the overall density of the system can be divided in a superfluid component and in a normal (i.e. non superfluid) one. The ratio of the averaged superfluid density and of the average total density is referred to as the superfluid fraction ρ_s, and is defined operatively through the ratio of the experimental I_e and of the classical I_c momentum

of inertia of a sample:

$$\rho_s = 1 - \frac{I_e}{I_c}. \tag{1.2}$$

In a completely superfluid sample $I_e = 0$ and $\rho_s = 1$. Superfluidity is accompanied by delocalization of particles into the sample, a condition referred to as off-diagonal long range order. More formally it is required that the single particle density matrix of the system $\rho(\mathbf{r}, \mathbf{r}')$ has the behavior:

$$\lim_{|\mathbf{r}-\mathbf{r}'|\to\infty} \rho(\mathbf{r}, \mathbf{r}') = n_0, \tag{1.3}$$

where $n_0 \neq 0$ is a constant. Of course the presence of this order in a solid is rather counter-intuitive, as is seems impossible to have delocalization occurring for particles pinned to lattice sites.

In recent years, the attention of theorists and experimenters alike has focused on solid ^{4}He as a potential supersolid system, following the observation of non-classical rotational inertia by Kim and Chan [6]. At the present time, agreement is still lacking, as to whether experimental findings indeed mark the first observation of supersolid behavior [7]. The most reliable theoretical studies, based on first-principle numerical simulations, show that superfluidity, if it occurs at all in solid Helium, is not underlain by the mechanism originally envisioned in the seminal works by Andreev, Lifshitz and Chester, i.e., through Bose Condensation of a dilute gas of vacancies or interstitials, [2, 5] but involves instead extended defects, such as dislocations [8]. In particular, a dilute gas of point defects in solid Helium has been predicted to be thermodynamically unstable [9]. Recently, new experiments seem to have resolved the controversy in favor of the non supersolid nature or the observed phenomenology [10].

Thus, it seems fair to state that solid ^{4}He does not afford a direct, simple, and clear observation of the supersolid phenomenon. Still, among all simple atomic or molecular condensed matter systems, Helium should be the best candidate by far, due to the favorable combination of large quantum delocalization of its constituent (Bose) particles, and weakness of the interatomic potential, therefore one might speculate whether the experimental observation of such phase will be possible at all. But what exactly, in the physics of this simple crystal, contributes to suppress (if not eliminate entirely) its superfluid response?

The thermodynamics of solid ^{4}He, as it emerges from first-principle quantum simulations, is largely determined by the strong repulsive core of the pair-wise inter-

atomic potential at short distance. For example, a very simple model of Bose hard spheres reproduces surprisingly accurately the phase diagram of condensed Helium. Such a repulsive core is a ubiquitous feature of ordinary interactions between atoms or molecules, arising from the Pauli exclusion principle, acting between electronic clouds of different atoms. One is then led to pose the theoretical question of which type of inter-particle interaction (or, class thereof) might underlie supersolid behavior. In particular, can an interaction featuring a "softer" core, saturating at short distance to a value of the order of the characteristic zero-point kinetic energy of the particles, result in the appearance of a supersolid phase?

This question might have seemed little more than "academic" until not so long ago, for how would one go about creating artificially such an interaction, which does not occur in any known naturally occurring quantum many-body system? However, impressive advances in cold atom physics appear to allow one to do just that, namely to "fashion" artificial inter-particle potentials, not arising in any known condensed matter system [11]. It makes therefore sense to search theoretically for supersolid, or other exotic phase of matters, based on more general types of interactions among elementary constituents than the ones considered so far, with the realistic expectation that such interactions might be realizable in the laboratory.

In this chapter, we study by computer simulations a system of Bosons in two dimensions interacting with a step potential, representative of a class of interactions for which the supersolid phenomenology occurs [12]. We employ the worm algorithm, an extension of the path integral Monte Carlo (PIMC), a finite temperature approach able to retrieve unbiased results for equilibrium thermal averages. We start by briefly reviewing the state of the current understanding of the physics of soft core particles, then results for the soft disk system are illustrated [13–15].

1.2 Soft core potentials in classical physics

Soft core potentials have been already intensively studied in classical physics. These were used as an effective potential for soft matters particles, such as polymers or macromolecules. To study the thermodynamic properties of a large assembly of such objects, one is most likely forced to disregard the internal degrees of freedom, as they are too many for a microscopic computation approach. Therefore an effective potential accounting for the interaction among these macro particles is needed. The interaction among these objects is rather different from the one occurring among their atomic constituents, which typically have a strongly repulsive core. For example,

in the case of low density objects such as macromolecules of low inner monomer concentration, it is possible for them to interpenetrate. While of course the inner components of the molecule do not overlap, two of these macromolecules may have the center of mass coinciding in space. Thus, the effective potential among these objects turns out to be soft core, being bounded to a finite value at zero separation.

This led to the study of the phase diagram of particles interacting through soft core interactions. In this context the discovery of multiple occupancy crystal occurred, a rather counter-intuitive phase in which a purely repulsive inter-potential can lead to particle "piling up" into the same lattice sites [16]. More formally, these are crystal lattices with a trivial basis of K particles in which all K basis vectors are null. The reason for this phenomenology is purely due to the potential energy gain upon clustering and is most easily understood in a simple example, see Figure 1.1. If one imagines a one dimensional crystal of particles interacting through a Gaussian two-body potential, for any density an extra particle added to the system will feel a potential having minima in the interstitial positions: no clustering therefore occurs. If instead we consider a system of particles interacting through a potential $V(r) \propto exp(-r^4)$ there exists a density above which the extra particle will feel a potential such that the lower potential region is on a lattice site rather than an interstitial position: this leads to clustering at high density and low temperature. In particular, a criterion is known within mean-field for the prediction of clustering, based on the form of the molecular interatomic potential [17]: its Fourier transform must have a negative value for some wave-vector $\mathbf{k}$. A system with this characteristic possesses in its ground state a cascade of cluster phases with progressively higher K upon increasing the chemical potential, separated among them by a first order phase transitions. Quite recently, a classical study of the phononic modes of a three dimensional cluster crystal was performed [18]: it was found that, besides the acoustic modes, the $3(K-1)$ optical modes of such a system are all degenerate and k-independent. It is reasonable to expect that the basic physics should remain relevant for a quantum-mechanical system as well. We expect clustering to occur also in the quantum case, but we will see that quantum fluctuations and Bose exchange will also give rise to completely new physics, including excitation modes not present in the classical case. From Figure 1.1 it is rather clear that clustering is associated to the existence of a potential barrier separating lattice sites. We anticipate that the height of this barrier will be important in the quantum case in determining the superfluid properties of the system.

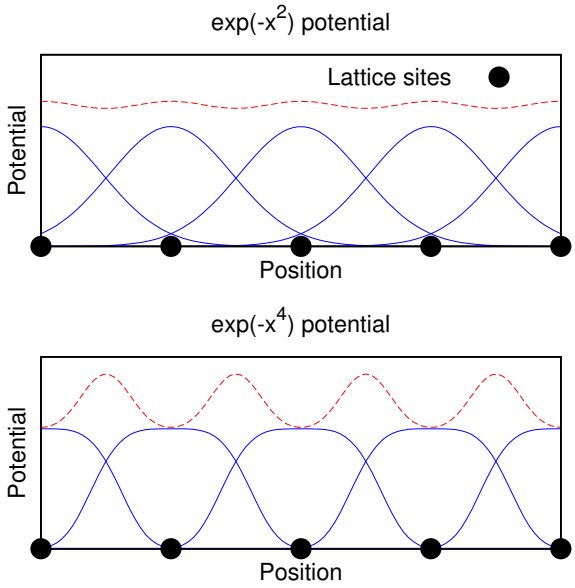

Figure 1.1: Individual (full blue) and total (dashed red) potential felt by a "test" particle in a one dimensional crystal of particles interacting via a Gaussian (top panel) and an $exp(-x^4)$ (bottom panel) potential. No clustering occurs in the Gaussian system at any density, while the lower system will display clustering instability at low temperature above some density.

1.3 Soft core Bosons

In Ref. [12], assemblies of Bosons in two dimensions interacting via soft-core potentials were studied by computer simulations. It was found that the system in its ground state displays clustering at high density, as expected in analogy with the classical case. Moreover, it was noted that near the liquid-cluster solid phase boundary, the crystal shows a finite superfluid fraction. This is of course the aforementioned supersolid phase. The potentials employed in [12] included long range tails going like $1/r^3$ or $1/r^6$, as they were representative of a class of potentials that may be fashioned in ultracold atom assemblies, via a Rydberg dressing mechanism [11]. However, a long-range tail is not required in order for a cluster crystal phase to exist. For example, the classical ground state of a system of particles interacting via the following, soft core potential

$$v(r) = \begin{cases} D & \text{if } r \leq a \\ 0 & \text{if } r > a \end{cases} \tag{1.4}$$

will be a cluster crystal at densities for which the mean inter-particle distance is less than the soft-core diameter a. However, it is not clear what role, if any, the long-range repulsive tail of the interaction plays in the occurrence of superfluidity of the cluster crystal. In order to identify the "minimal model" of supersolidity, we have investigated the low temperature properties of a two-dimensional system of spin zero Bose "soft disks", i.e., particles interacting via the simple potential given by Eq. (1.4). In spite of its simplicity, to our knowledge (and surprise) this has not been the subject of any prior theoretical study. The Hamiltonian of the system in reduced units is

$$\mathcal{H} = -\frac{1}{2}\sum_{i=1}^{N} \bigtriangledown_i^2 + D\sum_{i>j} \Theta(|1 - r_{ij}|), \tag{1.5}$$

where r_{ij} is the distance between particles i and j, all lengths are expressed in units of the soft-core diameter a, and all energies are expressed in units of $\epsilon_\circ = \hbar^2/ma^2$. The parameter $D \equiv V/\epsilon_\circ$ can also be expressed as $(a/\xi)^2$, where ξ is the quantum-mechanical penetration length of a potential barrier of height V. In the following we will sometime refer to particles which are interacting (i.e. are at a distance smaller than a) as being "overlapping". In the limit $\xi \to 0$ the model (1.5) reduces to the hard-sphere gas. The system is enclosed in a cell with periodic boundary conditions, of sides (L_x, L_y). We denote by N the average number of particles and express the density ρ in terms of the dimensionless parameter $r_s = 1/\sqrt{\rho a^2}$. In the crystal phase, we denote with K the average occupation per cluster. In order to study the system described by this Hamiltonian we use the worm algorithm approach in the grand canonical ensemble. This a finite-temperature stochastic technique, originating from the path integral Monte Carlo (PIMC) approach, able to give unbiased estimates of thermal properties at equilibrium. It is based on the path integral representation of the thermal density matrix of the system, where quantum particles are mapped into a classical system of interacting polymers. This technique is exact, meaning that all systematic and statistical errors can be in principle driven to zero. Details on the employed algorithms are described in appendix A.

1.4 Ground state Phase diagram

In this section we are in the low temperature limit; that is the value of the temperature is sufficiently low that most estimates can be regarded as essentially ground state

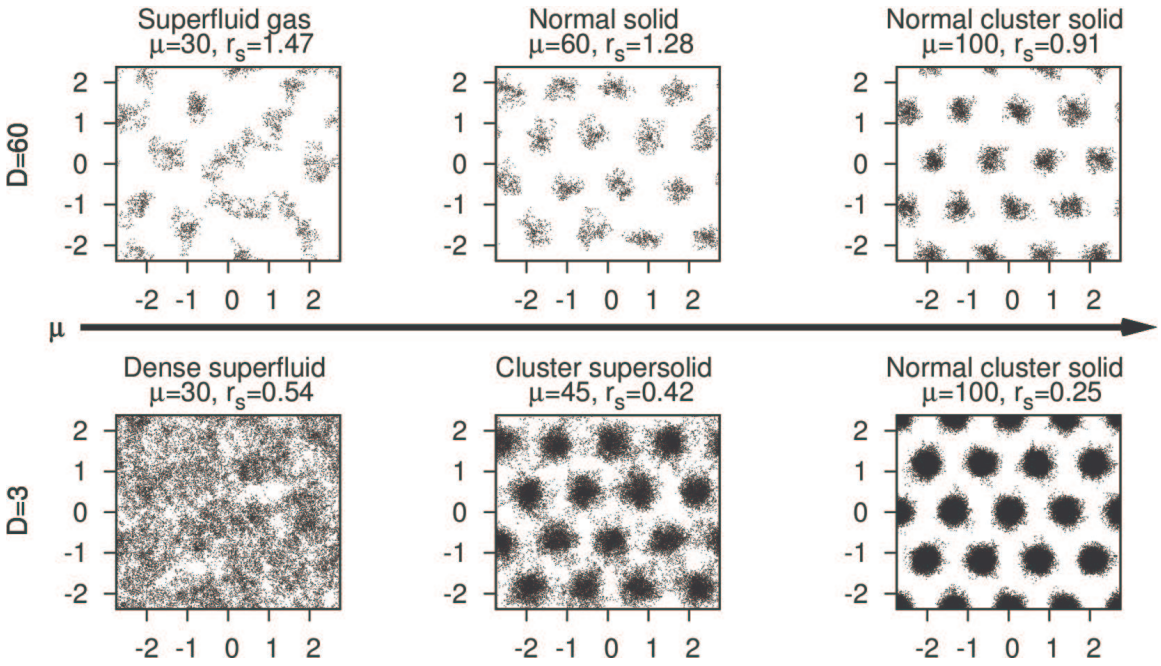

Figure 1.2: Qualitative low temperature phase diagram for high and low D as a function of μ. The panels show typical spatial configuration of the paths to which quantum particles are mapped, referring to the various phases. Results shown in the upper part of the figure correspond to simulations with D=60, whereas the lower part to D=3.

ones (typically $T \sim \epsilon_\circ$, for most quantities). We present here results obtained varying the chemical potential μ and D. Figure 1.2 summarizes qualitatively the ground state phase diagram in the $D - \mu$ plane, where each panels shows a snapshot of the paths representing quantum particles in each simulation box during the simulation. For $D \gg 1$ ($D = 60$ in the figure) and at low density, the physics of the system is essentially that of the hard-disk fluid, as particles tend not to overlap. In this condition the system is a superfluid gas. Upon increasing the chemical potential μ, the system undergoes solidification into a triangular crystal whit $K = 1$; across this structural transition the system loses completely the superfluid properties, therefore the system goes into a "normal solid" state. If the chemical potential is raised high enough ($\mu \approx D$), the system will no longer behave as being hard-disk and particles bunch into clusters (also referred to as "droplets") which organize in a solid preserving the triangular structure (a "cluster solid"). As mentioned above, the appearance at high density of such a cluster phase is a classical effect, directly related to the finite energy cost associated to particles being at a distance less than the soft core diameter. Indeed, by minimizing the potential energy per particle it is possible to estimate the

number of particles per cluster at a given nominal inter-particle distance r_s. Cluster formation becomes favorable for $r_s \lesssim 1$, and one finds $K = \alpha/r_s^2$, where α is a number $\lesssim 2$, *independent* of D. As in the classical case, in a cluster solid the lattice vector of the system is almost independent of the density, which affects mainly K.

For low values of D, the system can be kept superfluid even at high density ($r_s < 1$, meaning that particles do overlap). This is easily understood, as for weak interactions the potential energy gain due to crystallization becomes relevant only at high densities. Above a certain μ, the exact value depending on D, the system crystallizes directly into a cluster crystal with $K > 1$, as single occupancy crystal would be energetically unfavored. The interesting point is that in the cluster crystal phase near the liquid transition, even though density modulations appear in the system and translational symmetry is broken, the superfluid fraction ρ_s in the system is not fully suppressed: this is the supersolid phase. The presence of a global superfluid response in the supersolid phase, extending to the whole system, is assessed numerically through the direct computation of the winding number, Eq. A.28. This phase is completely analog to the one described in Ref. [12]; our result then shows that the long range tail included in the previous study is not a necessary condition for stabilizing the supersolid phase. The delocalization needed for the dissipationless flow to occur is determined by particle tunneling between neighbouring clusters. This is possible because near the transition the potential barrier present across clusters is relatively weak. Therefore, long cycles of exchanges of particles permutations between neighbouring lattice sites are allowed, eventually leading to the superfluidity of the system as a whole. Inside a cluster particles are at high density and essentially not interacting, as the soft disk potential is completely flat in a region of diameter 1, therefore it is reasonable to expect that individual droplets are Bose condensed (even though we cannot rigorously speak of Bose condensation in a two dimensional, confined system). In this condition the system behaves like a self-assebled lattice of superfluid regions connected by Josephson junctions. The kind of supersolidity we observe here is not related to Bose Condensation of lattice defects, but is rather of mean-field type, as similar behaviors are obtained within the Gross-Pitaevskii approach [11].

The character of the superfluid-supersolid phase transition is depicted in Figure 1.3, where results for $D = 5$, $(L_x, L_y) = (11.855, 10.267)$, 64 clusters, are reported. The static structure factor $S(\mathbf{k})$, order parameter for the breaking of translational symmetry, has a clear jump at the transition for $\mathbf{k}$ corresponding to a reciprocal lattice vector ($\mathbf{k} = [0.866\frac{2\pi}{a}, 0]$), taking place above $\mu = 38$, while the superfluid fractions drops from 1 to around 0.4. Immediately after the cluster transition $r_s = 0.59$ and

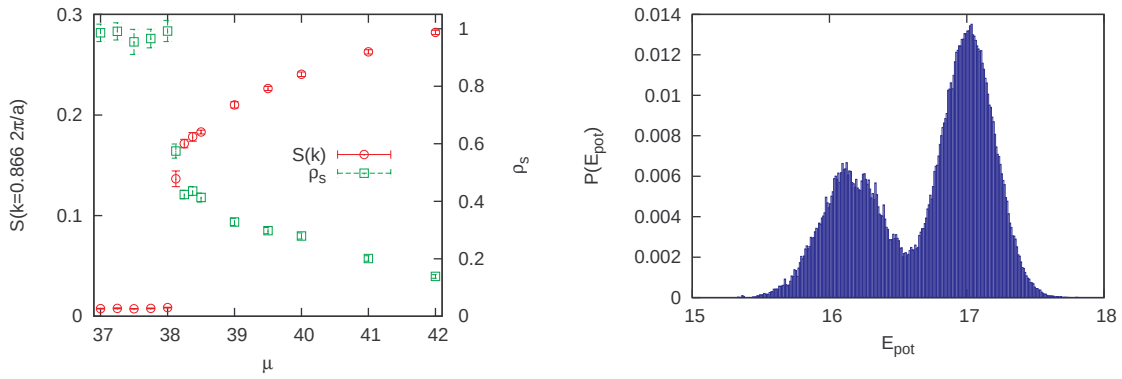

Figure 1.3: Static structure factor $S(\mathbf{k})$ ($\mathbf{k} = [0.866\frac{2\pi}{a}, 0]$ being a reciprocal lattice vector) and superfluid fraction ρ_s across the superfluid-supersolid transition (left panel), and distribution of the potential energy per particle E_{pot} estimator from the simulation near the superfluid transition ($\mu = 38.125$). Data are for $D = 5$, $(L_x, L_y) = (11.855, 10.267)$, 64 clusters. The jump in the $S(\mathbf{k})$ and the bimodal character of E_{pot} are a clear indication of a first order phase transition.

$K = 5.5$. The superfluid fraction is rather strong and not simply due to finite size effect, as we have verified that the value of 0.4 is already near the infinite size limit. Finally, near the transition, many estimators from the simulations are bimodal in character, indicating an oscillation between two thermodynamically stable states, as shown in Figure 1.3, right panel, for the potential energy per particle E_{pot}. Therefore we can unambiguously conclude that the superfluid-supersolid transition is first-order.

Increasing the density from the supersolid phase, ρ_s is progressively suppressed by the rising potential barrier between neighbouring lattice sites, which is associated with tunneling probability. This eventually destroys phase coherence in the system, which goes into a normal (or insulating, i.e. non superfluid) cluster phase. The same effect of tunneling suppression can be obtained also by increasing D at constant r_s, instead of μ. Higher potential barriers leads to increasingly compact clusters, i.e. particles pile up on a smaller spatial region (compare, e.g., the spread of the clusters in the two right panels of Figure 1.2).

The formation of the cluster crystal is also clearly reflected in the pair correlation function $g(r)$, as shown in Figure 1.4) for the case $D = 3$. Data in this Figure correspond to the lower panels of Figure 1.2. The pair correlation function is essentially featureless in the case of the dense superfluid, i.e., $g(r) \sim 1$ with only a slight depression for $r \approx 1$. For the supersolid, $g(r)$ develops a peak at $r = 0$, signaling multiple occupation of a single unit cell, as well as robust oscillations, with period consistent with the lattice constant of the cluster crystal. However in this phase particles tunnel

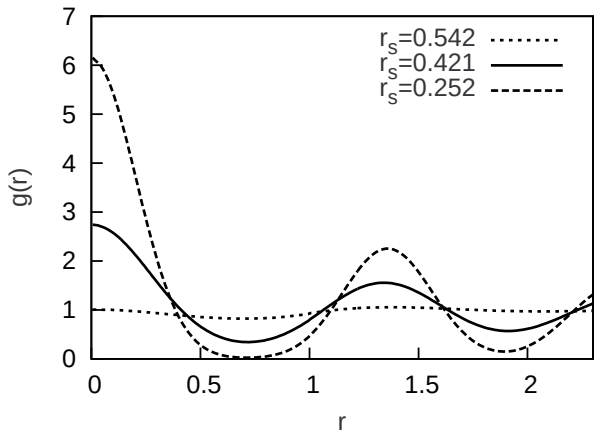

Figure 1.4: Ground state pair correlation function for $D = 5$, at three different values of r_s. The dotted line ($r_s = 0.542$) corresponds to a superfluid gas, the solid one ($r_s = 0.421$) to a supersolid, and the dashed one ($r_s = 0.252$) to a non-superfluid cluster crystal.

between adjacent clusters, therefore the pair correlation function takes on a finite value between successive peaks. As the density increases, clusters comprise a larger number of particles, as reflected by the increased strength of the peak at $r = 0$, and a depletion of the inter-cluster space occurs, in accordance with the increase of the inter-cluster barrier height. Similar effects are clearly depicted in Figure 1.5, where the averaged potential felt by a "test" particle is plotted along the crystallographic direction $[1, 0]$ passing trough the lattice sites for the supersolid and the cluster solid phases, for $D = 3$ at different densities. The supersolid retains a finite density of particles among neighbouring lattice sites, while in the cluster solid this is completely suppressed. Note that the potential felt by the "test" particle, which has minima in correspondence of the clusters, in the insulating phase is much deeper, resulting in the already noted stronger confinement of the droplets. As particles pertaining to the same droplet are essentially not interacting, the confining potential is originated exclusively by the nearest neighbouring clusters. Therefore the potential perceived by the "test" particle is precisely the cluster recoil force which determines phononic properties of the system (see forthcoming sections for excitation properties).

We conclude this section with a small outlook. We have depicted, in Figure 1.2, what is the behavior of the system in the limit of strong or weak D. However, it is not clear what is the behavior in the intermediate region. In particular it would be interesting to determine which is the minimum number of particles per unit cell K such that the supersolid phenomenology occurs, which would answer the question

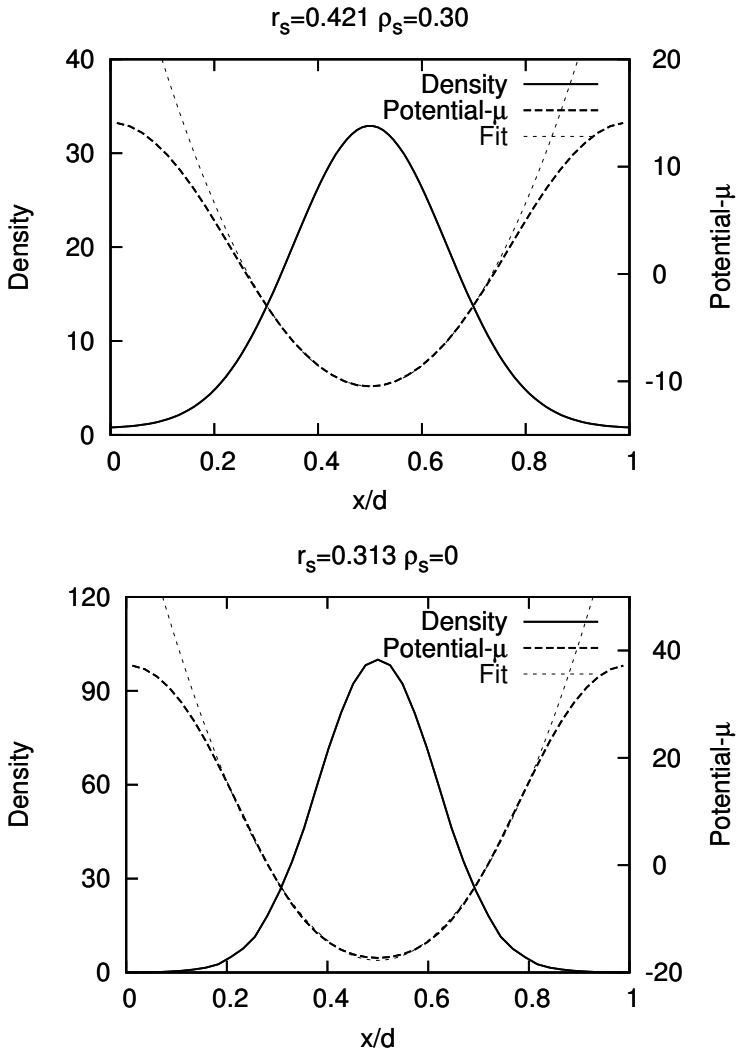

Figure 1.5: Density profile and potential felt by a test particle between two adjacent sites, for a supersolid conditions (upper panel) and for a cluster solid (lower panel), for $D = 3$ at different densities.

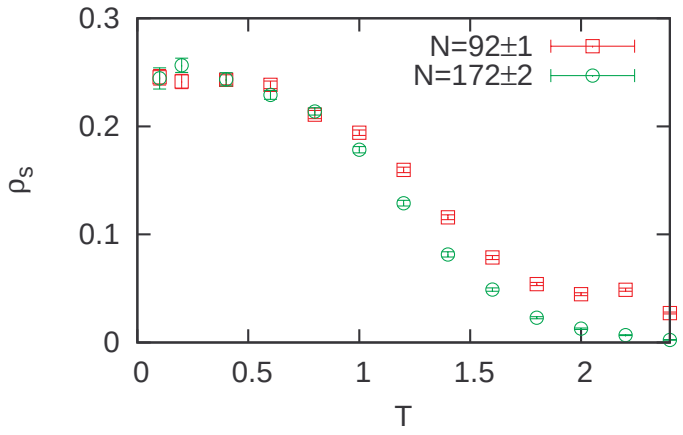

Figure 1.6: Superfluid fraction of as a function of T (in units of $\epsilon_\circ$) for a system of soft disks with D=5. Here, μ is set so that $r_s = 0.5$. Data shown are for two system sizes, comprising N=92 (circles) and N=172 (squares) particles.

as to if a supersolid with $K = 1$ can exists. Although from the test performed this seems not to be the case for soft disk Bosons, whether this is at all possible with some other potential remains unclear and is a potential subject for further investigations on similar systems.

1.5 Temperature dependence

A typical result for the dependence of ρ_s on the temperature in the supersolid phase is shown in Figure 1.6, for $D = 5$ and $r_s = 0.5$. As expected, the increase in temperature of the system results in the progressive suppression of the superfluid response. As in any simulation study, the transition is smeared by finite-size effects, and accurate finite-size scaling analysis of the result obtained on systems comprising significantly different numbers of particles would be required, in order to determine accurately the transition temperature. However, it can be seen that the $T = 0$ superfluid fraction is already near the thermodynamic limit, being equal within errorbars for the two different system sizes shown. The results shown in Figure 1.6 are consistent with a superfluid transition in the Kosterlitz-Thouless universality class, [19] as expected for a two-dimensional system. It is worth noting that the superfluid fraction does *not* saturate to a value of 1 as $T \to 0$. This is in line with the spontaneous breaking of translational invariance associated to crystalline order, as first pointed out by Leggett. [4]. The onset of superfluidity is well known to be underlain by long cycles of exchanges of identical particles (permutations). Figure 1.7 shows the fre-

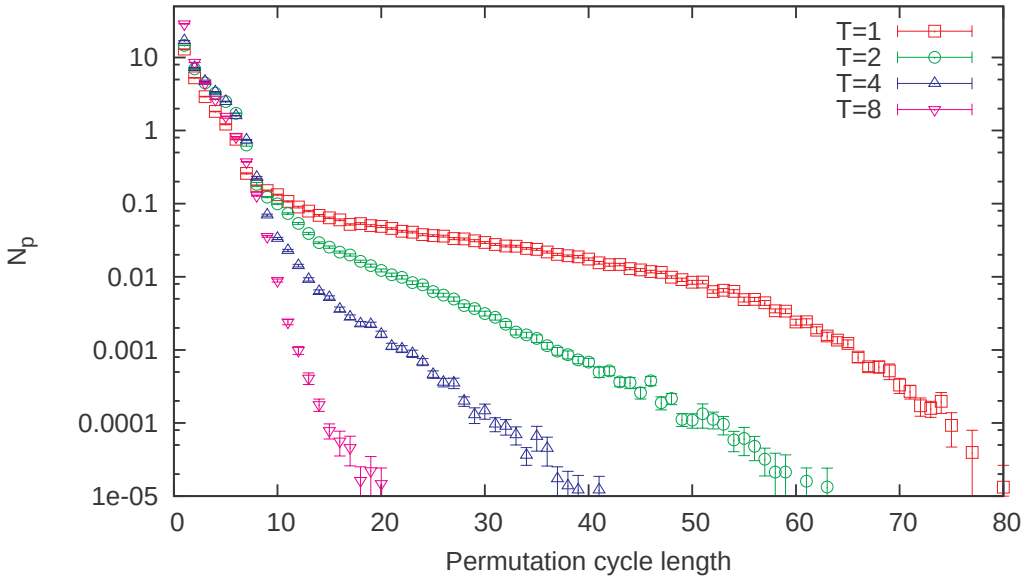

Figure 1.7: Number N_p of occurring permutation cycles in the simulation as a function of their length (i.e., the numbers of particles involved), at various temperatures. Temperatures are expressed in units of $\epsilon_\circ$. Here, $D = 5$ and $r_s = 0.5$. The system is in the cluster crystal phase at all temperatures.

quency of occurrence in the simulations of exchange cycles of different "length", i.e., involving different numbers of particles, at four different temperatures for a system in the cluster crystal phase. At the lowest T, exchanges involving all particles in the system are present. At the highest T, the number of occurring permutation cycles N_p drops sharply for cycles involving more than the number of particles per cluster (approximately 7 in the simulations shown here). However, even at the highest temperature represented here clusters are individually superfluid, even though global phase coherence does not exist.

Regarding the transition to normal cluster crystal of the supersolid upon increasing the density, our data at finite T are consistent with an exponential decrease with D or μ of the superfluid transition temperature of the cluster crystal, but we cannot rule out continuous quantum phase transitions between a supersolid and a normal cluster crystal, driven by either μ or D. In conclusion, we can state that the supersolid phase, at constant D, can be suppressed either by increasing the temperature or the density, in which case the superfluid character is lost, or by lowering the density, in which case it is the crystalline structure to be suppressed. The resulting schematic finite temperature phase diagram in the limit of small D, compatible with data from our simulations, is shown in Figure 1.8.

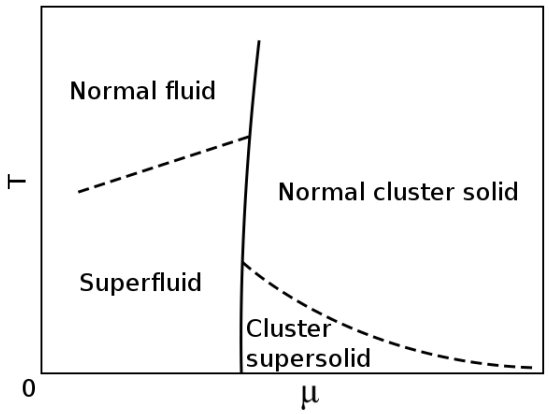

Figure 1.8: Qualitative low D phase diagram at finite T. Thick lines show first-order, dashed lines continuous phase transitions.

1.6 Momentum distribution

The finite superfluid response is accompanied by a strong signature in the momentum distribution $n(\mathbf{k})$, shown for low temperatures in Figure 1.9, along the crystallographic direction $[0, 1]$. Even though Bose condensation is suppressed in two dimensions at finite temperature, in the supersolid phase the calculated momentum distribution features a pronounced peak at $k = 0$, as well as a smaller peak at $k = 2\pi/l$. Neither is found in the insulating crystal. The former represents the integral of the occupation number of low-momentum states, corresponding to a slow power-low decay of the one-body density matrix (OBDM). The latter describes a substantial real-space modulation of the OBDM with the periodicity of the triangular lattice.

1.7 Leggett Bound

It is interesting to consider the upper limit on the superfluid fraction induced by density modulations, as pointed out by Leggett [4]. The bound reads as [20]:

$$\begin{aligned} \rho_s \leq \min_{\varphi(\mathbf{r})} \left(\frac{1}{V v_0^2} \int \frac{\rho(\mathbf{r})}{\rho} \mid \mathbf{v}_s(\mathbf{r}) \mid^2 \mathrm{d}\mathbf{r} \right) \\ \mathbf{v}_s(\mathbf{r}) = \left(\frac{\hbar}{m} \right) \bigtriangledown \varphi(\mathbf{r}) \\ \varphi(\mathbf{r} + (n\, L_x, n'\, L_y)) = \varphi(\mathbf{r}) + 2\pi n'', \end{aligned} \tag{1.6}$$

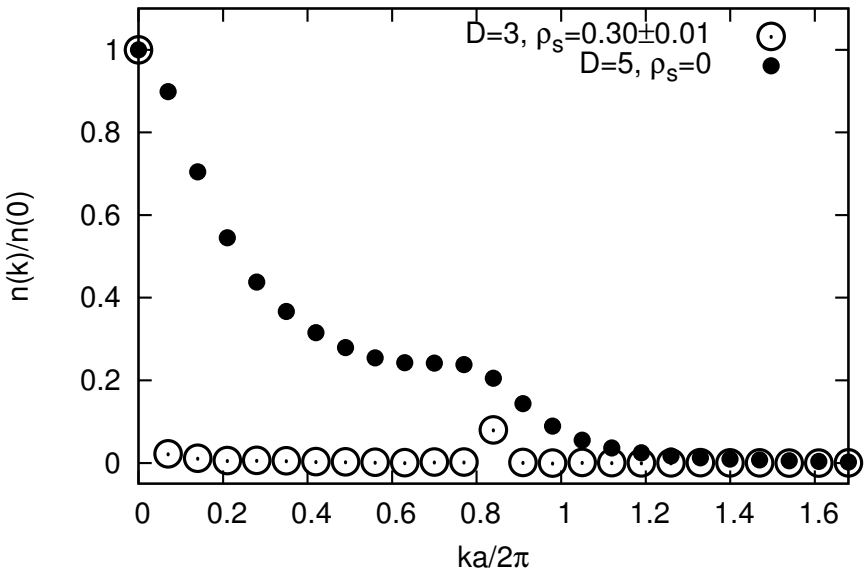

Figure 1.9: Ground state momentum distribution $\mathbf{k}$ for $r_s = 0.421, \mu = 45$ in the supersolid phase ($D = 3$, open circles) and in the insulating cluster solid ($D = 5$, full circles). The wave-vector spans the range 0–$\frac{4\pi}{d\sqrt{3}}$ along the direction $[0, 1]$. The peaks at $k = 0$ and $k \sim 0.85$ (in units of $2\pi/a$) are due to (quasi)condensation and strong modulation of the OBDM with the lattice periodicity, respectively.

where $\varphi(\mathbf{r})$ is a phase function, ρ and v_0 are the average value over the simulation cell of $\rho(\mathbf{r})$ and $\mathbf{v}_s(\mathbf{r})$ respectively, and n, n' and n'' are integers. This equation essentially states that if the system is uniform the superfluid fraction can be one, while if density modulations are present its value must be necessarily lower. Although the Leggett bound is very loose in strongly correlated systems such as solid ^{4}He or para-hydrogen [20], it could provide an argument to explain why soft disks are supersolid under appropriate conditions while two-dimensional solid ^{4}He is never, should their density profiles look different. For example, one may expect the density profile of a cluster supersolid to be less modulated than that of an ordinary condensed matter system. In order to verify this possibility we compared (Figure 1.10) the density profile of the supersolid soft disk system for D=5, T=0.1, μ=45, $\rho_s = 0.25 \pm 0.01$ with the one of solid ^{4}He in two dimensions at low temperature slightly above the melting density ($\rho = 0.0765\mathring{A}^{-1}$, $T = 1K$). Actually the two density profiles are remarkably similar, and the Leggett bound gives $\rho_s \leq 0.39$ for the soft disk and $\rho_s \leq 0.37$ for Helium. This shows that, within the constraint of Eq. 1.6 imposed by one-body properties, many-body effects have even qualitative effects on the superfluid response. In particular, the difference between hard- and soft-core systems is entirely

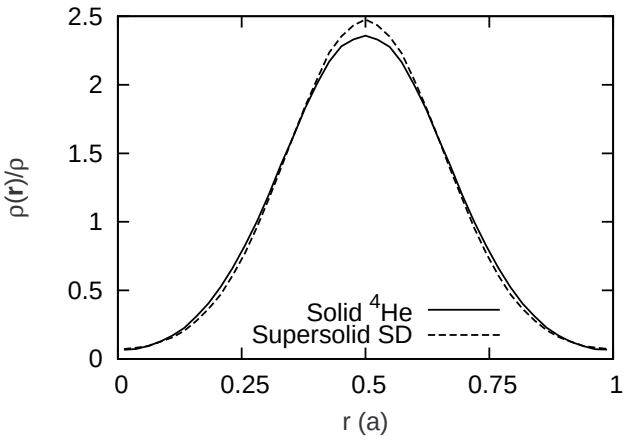

Figure 1.10: Density profiles $\rho(\mathbf{r})/\rho$ for soft disks in the supersolid phase (D=5, T=0.1, μ=45,$\rho_s = 0.25 \pm 0.01$), and for solid ^{4}He near melting ($\rho = 0.0765\mathring{A}^{-1}$, $T = 1K$). These density profiles are obtained along the crystallographic direction $[1, 0]$ passing trough a lattice site. r is in units of the lattice parameter.

encoded in two-body (and possibly higher) correlations. Indeed the pair distribution function of supersolid soft disks is peaked at contact, as opposed to the correlation hole induced by hard-core repulsion in Helium.

1.8 Toward realistic systems

If supersolid systems could be realized in the field of ultracold atoms, in principle it would be desirable to compare results from the experiments with Monte Carlo simulations, as it was already done for lattice systems [21]. In order to do so, one would need to include some features of realistic interparticle potentials. In particular, in ultracold atom systems a soft core potential may be realized employing Rydberg dressed atoms [11], which can produce long range (in the order of μm), soft-core potentials. However, the atoms still possess an hard core contact interaction, although of a typical size much smaller (order of nm). Therefore it is be interesting to check if this hard core interaction would break the phenomenology depicted in the previous sections. It is easy to convince ourselves that, for a classical system at zero temperature, if the excluded surface of this hard core is much smaller than the inverse density inside a droplet in the system, the clustering phenomenology should still be present. However, it is less obvious that the supersolid phase should survive to zero point motion when quantum fluctuations are present. Indeed, simulations in which a small excluded volume is added to Hamiltonian 1.5 display stable supersolid

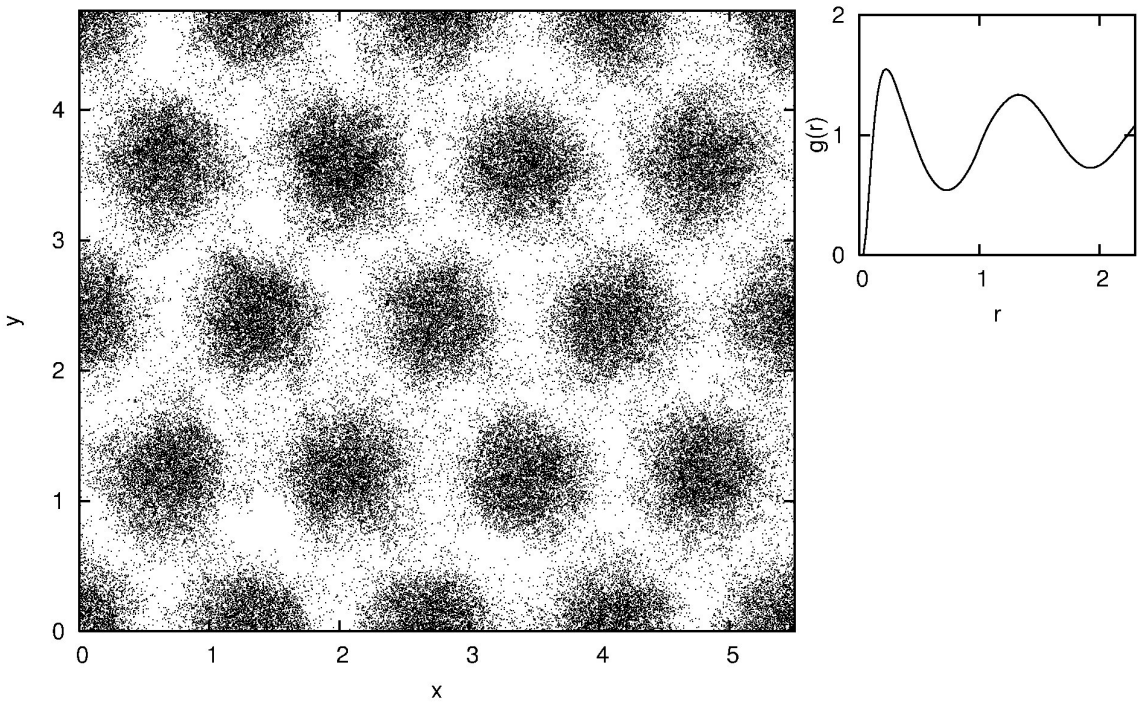

Figure 1.11: Snapshot of a Monte Carlo configuration for a soft disk system with hard core repulsion in the supersolid phase ($\rho_s = 0.27 \pm 0.04$). In the smaller panel the $g(r)$ of the system is reported, being zero at contact for this case.

phases for certain parameters values. In Figure 1.11, a configuration snapshot from a simulation of the supersolid system is reported, together with the pair distribution function which is now zero at contact. This simulation is performed in the canonical ensamble, for 144 particles ($r_s = 0.427$) having a small hard core of diameter $0.05a$, with the soft-core (diameter a) having $D = 17$. The resulting superfluid fraction is $\rho_s = 0.27 \pm 0.04$.

Moreover, the experimental realization of a supersolid system in the ultracold atom field would be inevitably trapped in an harmonic potential. In such a confining potential the supersolid system would be surrounded by a liquid region. It is interesting to calculate properties that can be compared directly with experiments, such as the momentum distribution $n(\mathbf{k})$. We give in Figure 1.12 a couple of examples of such simulations (for $D = 5$, $\mu = 50$, no contact interaction included), to illustrate the qualitative features expected in this situation. The supersolid phase (top panels) is indeed surrounded by a broad superfluid region. However, in the angle averaged $n(k)$ it is possible to observe the oscillations already noted in the extended system, although strongly smoothed. In a normal liquid (bottom panels), we recover the broad

distribution expected for a non condensed system.

1.9 Excitation spectrum

We now come back to the soft disk system (no hard-core interaction) in the extended case. Of particular interest is the spectrum of elementary excitations of a supersolid, which offers access to arguably even more cogent information on the physics of a system, than structural or energetic properties of the ground state. For example, it is not obvious how the excitation spectrum would combine specific traits of a solid and of a superfluid, i.e., whether two separate Goldstone modes should be present, reflecting the two broken symmetries, or a single mode of distinct, different character. Also of interest is establishing whether the excitation spectrum of a superfluid system, that also breaks translational invariance, displays a roton minimum. It is also worth noting that the experimental study of the dynamic structure factor in assemblies of ultracold atoms has recently begun [22, 23], therefore results presented here are in principle amenable of direct experimental verification. We use the Worm Algorithm in the continuous-space path-integral representation to simulate the system described by Eq. 1.5 in the grand canonical ensemble (i.e., at fixed temperature T, area A and chemical potential μ). The simulation gives an unbiased, accurate numerical estimate of the imaginary-time intermediate scattering function

$$F(\mathbf{k},\tau) = \langle \hat{\rho}_{\mathbf{k}}(\tau)\hat{\rho}_{-\mathbf{k}}(0)\rangle/N, \tag{1.7}$$

where $\hat{\rho}_{\mathbf{k}} = \sum_j e^{i\mathbf{k}\cdot\mathbf{r}_j}$ is the density fluctuation operator at wave-vector $\mathbf{k}$ and the brackets denote a thermal average. The dynamic structure factor $S(\mathbf{k},\omega)$, which measures the excitation spectrum of the density fluctuations, is related to $F(\mathbf{k},\tau)$ via an inverse Laplace transform:

$$F(\mathbf{k},\tau) = \int \mathrm{d}\omega e^{-\tau\omega} S(\mathbf{k},\omega). \tag{1.8}$$

It is well known that there exists no general scheme to invert a Laplace transform from noisy data, in a way that is reliable, accurate and controlled. However, for physical spectra whose dominant contribution is given by a few well-defined peaks, some techniques are able to identify satisfactorily locations and spectral weights of those peaks. In this work, we made use of the Genetic Inversion via Falsification of Theories (GIFT) [24] approach for the numerical inversion of the Laplace transform Eq. 1.8.

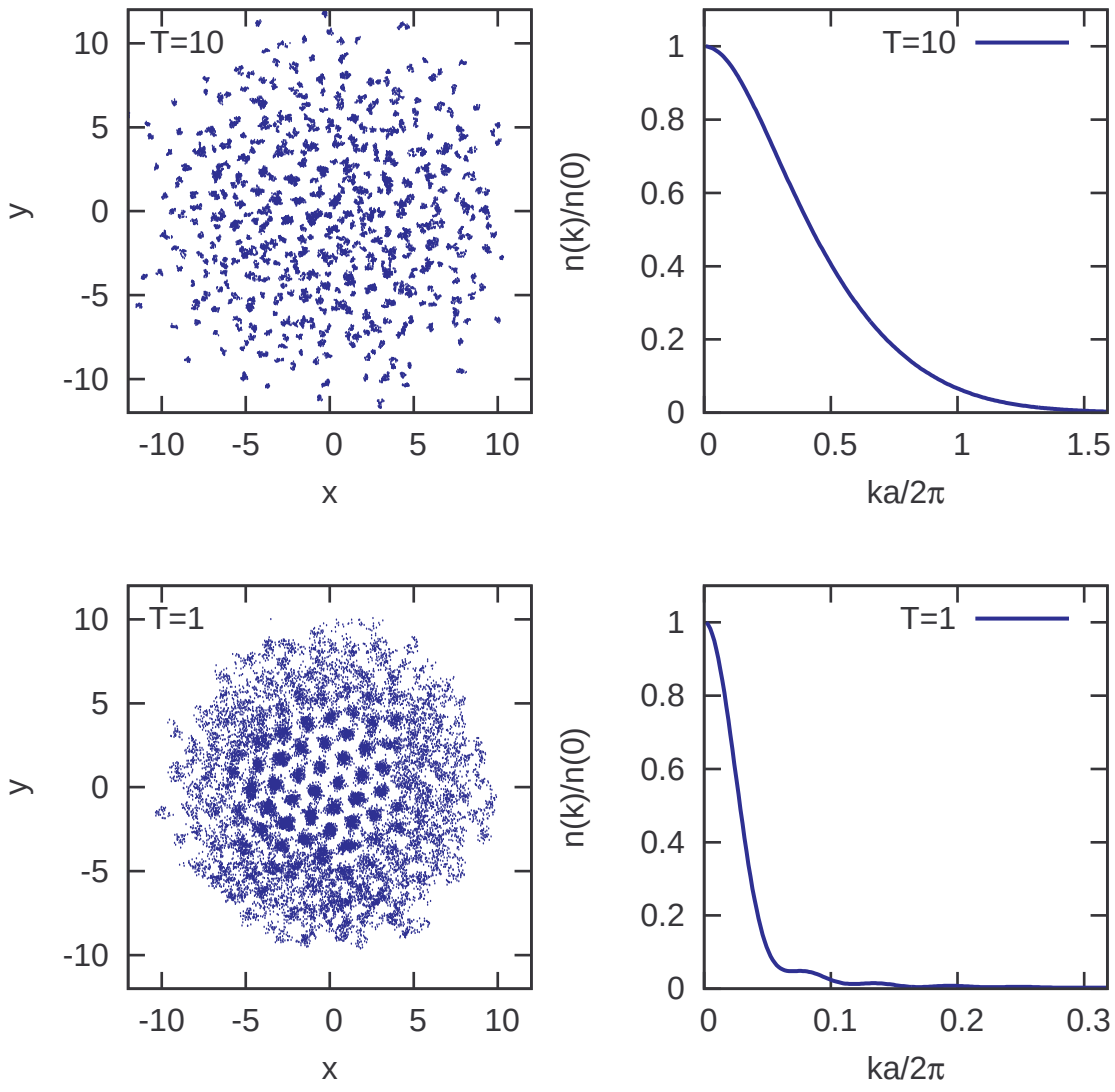

Figure 1.12: The cluster supersolid is stable even in a trapped systems, although surrounded by a superfluid region (low-left panel). In the top-left a non superfluid liquid configuration is reported. In the right column the angle averaged $n(k)$ are reported, with the supersolid showing a clear condensation peak with secondary peaks associated with crystallization of the system.

When applied to superfluid ^{4}He, the GIFT method has been shown to separate correctly the sharp quasiparticle peak of the phonon-maxon-roton elementary excitation from the broad multiphonon contribution [24], whereas the more commonly adopted Maximum Entropy scheme [25] tends to merge both structures [26, 27]. Alongside with this method, using the information on the number n of excitations visible in the reconstructed spectrum, we compute the energy of the n observed excitations by assuming the spectrum $S(\mathbf{k}, \omega)$ as a function of ω only is constituted by n delta functions (n-pole approximation) and fit their positions and strengths to the available $F(\mathbf{k}, \tau)$ data. Figure 1.13, colormap, shows GIFT reconstructions of the dynamical structure factor $S(\mathbf{k}, \omega)$ in the superfluid, supersolid and cluster crystal phases; datapoints are obtained from the n-pole approximation instead. In the superfluid phase (Fig. 1.13 a), the spectrum is characterized by the usual phonon-maxon-roton dispersion, with the notable peculiarity that the roton minimum is located just short of $2\pi/a$, rather than around $2\pi/r_s$. This suggests that the incipient crystallization takes place with a lattice parameter larger than the mean inter-particle distance. Indeed, upon increasing μ, the superfluid undergoes a first order phase transition into a triangular cluster (super)solid with a lattice spacing d somewhat larger than a [13].

The spectrum of the cluster crystal is also standard (we have studied longitudinal excitations only). Figure 1.13c) shows that, within the first Brillouin zone, most of the spectral weight is concentrated in an acoustic phonon band. We observe a non-negligible zero-frequency contribution at all wavelengths, representing a diffusive mode of lattice defects (phase-incoherent hopping of particles between multiply occupied sites). We also observe for some $\mathbf{k}$-points the presence of an optical mode, which might be the analogue of the highly degenerate breathing mode of individual clusters observed for classical cluster systems [18], but we do not determine its dispersion relation here.

The excitation spectrum of the supersolid phase (Fig. 1.13b), is the main result of this section. The spectral weight is clearly partitioned in two distinct branches. The higher-energy mode is a longitudinal acoustic phonon, with a linear dispersion at small $\mathbf{k}$ and near the reciprocal lattice vector (the end of the $\mathbf{k}$ scale in the figure), and frequencies between the phonon-maxon of the liquid and the longitudinal phonon of the solid. This assignment is further supported by the following analysis. We computed the average potential $v_{\text{test}}(r)$ felt by a test particle across a lattice site (upper panel of Figure 1.5, also showing the particle density profile). We define a force constant by fitting a quadratic potential to the bottom of $v_{\text{test}}(r)$, and obtain the phonon frequency of an harmonic crystal with that force constant and the aver-

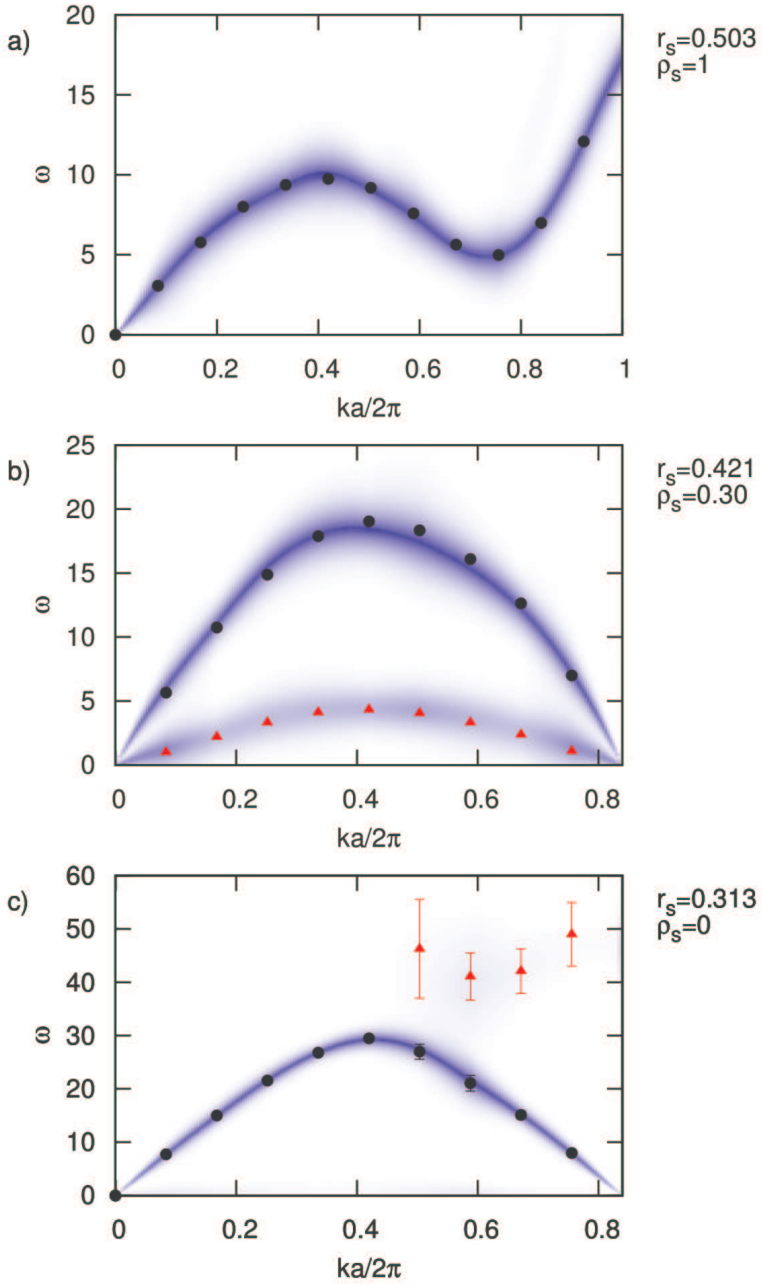

Figure 1.13: Dynamic structure factor $S(\mathbf{k}, \omega)$ in various phases. Panels refer to a) a superfluid, b) a supersolid, and c) a non-superfluid cluster crystal. The colormap is obtained by smoothing and interpolation of the calculated GIFT spectra. In order to emphasize the dispersion of the spectrum, for each $\mathbf{k}$ the dynamic structure factor is rescaled to a common maximum value in all panels. The datapoints are obtained from the n-pole approximation (see text); when errorbars are not reported, these are of the order of the symbol size or smaller. For the modulated phases b) and c), the primitive vectors of the Bravais lattice are $d(1,0)$ and $d(1/2, \sqrt{3}/2)$ with $d = 1.375a$; the wavevector spans the range 0–$\frac{4\pi}{d\sqrt{3}}$ along the direction $[0,1]$. The mean site occupation is $K = 9.2$ in b) and $K = 16.7$ in c). Also given are the values of r_s (mean inter-particle distance) and ρ_s (superfluid fraction). The upper (lower) panel corresponds to the supersolid (cluster solid) phase of Figure 1.5

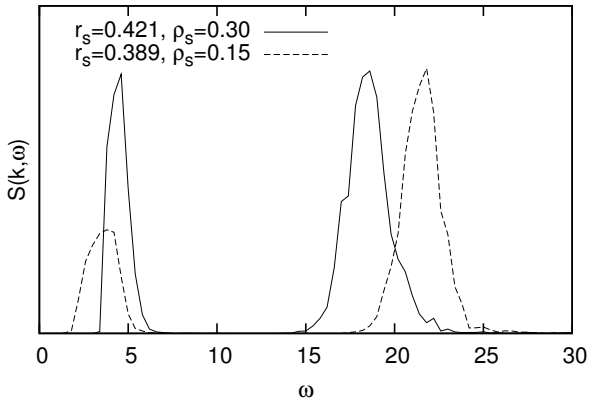

Figure 1.14: Comparison of $S(k,\omega)$ at different densities. They are calculated at the edge of the Brillouin zone, for $k = \frac{2\pi}{d\sqrt{3}}$. The solid (dashed) line pertains to a supersolid with $r_s = 0.421$ and $\rho_s = 0.30$ ($r_s = 0.398$ and $\rho_s = 0.15$).

age mass of a cluster, given by K. The sound velocity of the harmonic crystal is in satisfactory agreement (to within $\sim 10\%$) with the slope of the phononic branch of the supersolid displayed in Figure 1.13 b).

The lower branch of the supersolid spectrum is also acoustic. Its nature is assessed by studying its behavior as the superfluid fraction ρ_s decreases on approaching the transition to the normal cluster crystal. Figure 1.14 displays $S(k,\omega)$ at the Brillouin zone edge for $r_s = 0.421$, as in Figure 1.13b), and for a denser system at $r_s = 0.389$ where the system is still supersolid but ρ_s drops from 0.30 to 0.15. When the density increases the particles progressively get more localized around lattice sites, as shown by the density profiles of Figure 1.5, and the system gets stiffer. Correspondingly, the high energy peak of the supersolid spectrum shifts to higher frequencies, as expected for a phonon-like excitation. The low energy mode instead loses spectral weight following the loss of superfluid fraction, and shifts to lower frequencies, reducing its bandwidth. Similar results are obtained if superfluidity is suppressed by increasing the strength (D) of the interaction, at fixed density. The lower branch is thus seen to be largely unrelated to the spectra of the superfluid or the cluster solid phases. It could only be related to the phonon-maxon-roton of the superfluid if the density modulation of the supersolid could vanish smoothly. This possibility is preempted, however, by the first-order phase transition at the melting density of the supersolid.

The structural properties of the cluster crystal [13], with particles hopping between adjacent lattice sites, suggest an analogy with the Bose-Hubbard model which is clearly born by the spectral properties of both systems. While in continuous space

the lattice is self-assembled, and thus features vibrational modes that are obviously absent in the BHM, the lower branch of the supersolid phase shows noteworthy similarities with the acoustic excitation present in the superfluid phase of the BHM. In particular, its behavior for wavelengths approaching the reciprocal lattice vector $\frac{4\pi}{d\sqrt{3}}$ parallels the expected linear vanishing of the superfluid mode of the BHM. Our calculations use a finite simulation cell, implying a minimum distance of the wavevector from $\frac{4\pi}{d\sqrt{3}}$. Within this limitation, we find that the spectrum of the cluster supersolid is gapless at the reciprocal lattice vectors, i.e. there is no roton minimum. The analogy between the supersolid cluster crystal and the BHM superfluid is further supported by the softening of the low energy branch of the supersolid observed for increasing density (and/or increasing D), Figure 1.14. A higher density implies reduced hopping probability and enhanced on-site repulsion, as shown in Figure 1.5. In the BHM this is equivalent to lowering the t/U ratio, which in turn is known to reduce the bandwidth of the superfluid acoustic mode of the superfluid phase [28]. The phase transition between supersolid and cluster solid observed in the SD system has thus some bearing with the superfluid to Mott insulator (MI) quantum transition in the BHM. This is not a full correspondence, as finite temperature and lattice dynamics contribute to stabilize a compressible cluster solid phase with non-integer mean occupation of lattice sites, more similar to the normal liquid (NL) than the MI phase of the BHM; on the other hand, the finite temperature NL and MI phases are not fundamentally different, being connected by a crossover upon varying t/U [29].

In the BHM, where the translation invariance is explicitly broken, the presence of an acoustic mode is due to long-range phase coherence, which breaks the continuous gauge symmetry. In a continuous space superfluid, the corresponding mode is second sound. This mode is not seen in the spectrum of the SD superfluid phase (Figure 1.13a), where its spectral weight is presumably exceedingly low. In view of the analogy between the SD supersolid and the BHM superfluid, we are led to suggest that the lower branch of the supersolid is a kind of second sound. Indeed, the presence of a second longitudinal acoustic branch in a supersolid is a common feature of several phenomenological models, which assume some degree of phase coherence and the possibility of a density modulation not commensurate to the particle number [30–32]. In particular, Ref. [30] characterizes this second branch in the supersolid phase as a Brillouin peak, with out-of-phase fluctuations of the normal and the superfluid density, emerging from the defects-associated Rayleigh mode of the normal solid, much as second sound in superfluid ^{4}He [33].

The experimental realization of a SS system similar to that studied here appears

possible in assemblies of ultracold atoms [12]. The double acoustic excitations, peculiar to the supersolid phase, could be detected via Bragg Spectroscopy [22, 23]. This could be an interesting experimental verification of our findings, and a tool for identifying the supersolid phase unambiguously as well.

1.10 Conclusions

In summary, we have studied by Monte Carlo simulations a two-dimensional system of Bosons on the continuum interacting via a repulsive, short-range soft-core potential. This system displays a low-temperature supersolid phase, wherein particles tunnel across nearest-neighbouring, multiply occupied unit cells. We studied the qualitative phase diagram of the system, characterizing the conditions under which the supersolid phase can be observed. We analyzed various properties of the system in its ground state, such as the pair correlation function and momentum distribution. Through the computation of the Leggett bound it was shown that the superfluidity in the system is not due to a shallower density profile with respect to solid ^{4}He, but is encoded in higher order correlations. The dependence of the supersolidity phenomenology on the temperature was found to be compatible with a Kosterlitz-Thouless behavior. We also carried out a numerical study of the excitation spectrum of the model. The main finding is that two well-defined, distinct acoustic modes are present in the supersolid phase. The higher-energy branch is determined by the lattice dynamics, while the softer mode is uniquely due to the presence of a finite superfluid fraction. Its dispersion closely parallels that of the excitation spectrum of a superfluid Bose-Hubbard model: it further softens as superfluidity is demoted approaching the insulating solid phase; furthermore, for the system sizes studied here, it looks linearly vanishing at the reciprocal lattice vectors, rather than featuring a finite roton minimum. The physics illustrated here should be observable under relatively broad conditions, if soft-core pair-wise interaction potentials could be fashioned.

Chapter 2

Minimum Energy Pathway via Quantum Monte Carlo

2.1 Introduction

Determining minimum energy pathways (MEP) of reactions is of fundamental importance in scientific and technological applications. The knowledge of barrier heights is key to the prediction of catalytic properties of materials since it enables the use of transition state theory (TST) to determine reaction rates [34–36]. Locating efficiently the transition state on a potential energy surface (PES) is in fact a popular subject in computational physics and, to this aim, a variety of algorithms have been developed such as the shallowest ascent, synchronous transit, and nudged elastic band (NEB) approaches [37]. All these techniques ultimately rely on a method to determine the energy of a given atomic configuration and/or its derivatives.

If we restrict ourself to quantum simulations, the most used approaches are density functional theory (DFT) or highly-correlated quantum chemical methods, that is, wave function post-Hartree-Fock techniques such as the coupled cluster single-double and perturbative triple approach (CCSD(T)), which is generally considered the "gold standard" in quantum chemistry. Many of these wave function methods are variational (though coupled cluster methods are not) and in principle offer a systematic route to converge toward the exact energy, even though the increasing computational cost and the slow convergence severely limits this possibility. Their main drawback is that all these approaches implicitly or explicitly rely on expanding the wave function in Slater determinants and, therefore, require large amount of computer memory and have a poor size scaling (N^7 for CCSD(T), N being the number of electrons), limiting

their range of applicability to small systems.

Consequently, for larger systems DFT remains the method of choice due to its much more favorable computational cost (scaling from N^2 to N^4). Even though continuous progress in the field has led to the development of more precise and sophisticated DFT functionals, the situation is still far from satisfactory if one aims at high accuracy [38, 39]. For example, it is well known that popular functionals such as B3LYP [40–42] often lead to poor transition state geometries and barrier heights [43, 44]. Moreover, there remains a degree of arbitrariness in the choice of the functional, with different functional providing qualitatively different results in some cases. Since DFT methods are not variational and do not offer a systematic way to improve their estimates, one has to resort to different approaches if better accuracy is needed.

Alternatively, one can employ quantum Monte Carlo (QMC) methods, such as variational (VMC) and diffusion (DMC) Monte Carlo. These well-established ab-initio techniques take advantage of Monte Carlo integration over the full Hilbert space. In particular, VMC is a stochastic way of calculating expectation values of a complex trial wave function, which can be variationally optimized. DMC provides instead, with a higher computational cost, a stochastic ground-state solution to the full Schrödinger equation, given a fixed nodal surface (using the fixed-node approximation in order to avoid the notorious fermion sign problem). A more detailed description of these algorithms is provided in appendix B. Because integrations are performed in the full Hilbert space, one can make use of non separable wave functions, with the explicit electron-electron correlation encoded in a so-called Jastrow factor or in a backflow transformation [45]. This allows for noteworthy accuracy already using a simple and non memory-intensive single determinant Slater-Jastrow wave function. Although considerably more expensive than DFT methods (scaling as N^3 with a much larger prefactor), DMC generally offers better accuracy with respect to DFT [46–48], although some unsatisfactory cases are known [49–51]. Furthermore, QMC methods possess a variational principle, which is a useful feature when one has to evaluate energy differences as in TST. From a computational point of view, QMC codes can be made to scale linearly with the number of cores and are not particularly memory demanding, making them suitable for today's massively parallel supercomputers. Finally, QMC methods offer in principle the possibility to push the calculation up to a desired accuracy by employing wave functions of increasing complexity (although, from a practical point of view, one is likely to adopt simple wave functions for intermediate-to-large sized systems due to the increased computational cost of multi-

determinant wave functions). One of the alleged drawbacks of QMC methods is its presumed inability to calculate energy derivatives effectively. Although various authors have addressed the problem from different perspectives [52–55], the application of forces from QMC has been limited to few cases [56–58]. We will in this chapter make extensive use of forces with the approach described in [52], and we hope to convince the reader that QMC forces are actually reliable and computationally viable.

In previous QMC studies of reaction barriers the geometries have been taken either from DFT, or from constrained geometry optimization along an assumed reaction coordinate [58–60]. In the present work, we show nudged elastic band and climbing image calculations [61, 62], where the geometry optimization of all the NEB images is done fully at the QMC level. We believe that assessing the performance of QMC methods for a set of simple reactions is useful, as there is no literature on the determination of reaction paths within QMC and there are limited informations on its performances [58–60]. The QMC calculations are performed with a modified version of the CHAMP program, a quantum Monte Carlo program package written by C. J. Umrigar, C. Filippi, and collaborators. We also use the GAMESS package [63], for producing initial wavefunction and DFT calculations. For some representative challenging reactions from the NHTBH38/04 database [44, 64] and for a hydrogen transfer reaction [65], we determine transition state geometries and forward-reverse barrier heights within VMC and DMC, and compare our results against several current DFT functionals and other wave function methods. We demonstrate that VMC is able to locate reaction geometries with higher accuracy than DFT, while DMC outperforms DFT in evaluating barrier heights [66].

2.2 Methodology: Wavefunction and Pseudopotentials

In the following, unless otherwise stated, we employ a single determinant Slater-Jastrow wave function for our VMC/DMC calculations:

$$\Psi(\mathbf{r},\mathbf{R}) = D^{\uparrow}(\phi^{\uparrow}, r^{\uparrow}, \mathbf{R}) D^{\downarrow}(\phi^{\downarrow}, r^{\downarrow}, \mathbf{R}) J(\mathbf{r},\mathbf{R}), \tag{2.1}$$

where $\{\mathbf{r}\}$ and $\{\mathbf{R}\}$ denote the full set of single particle electronic ($\mathbf{r}_i$) and nuclear ($\mathbf{R}_i$) positions, respectively. The Jastrow factor, $J(\mathbf{r},\mathbf{R})$, explicitly depends on the

inter-particle distances, [67]:

$$J(\mathbf{r},\mathbf{R}) = exp\left\{-\sum_{i,j} U_{ee}\left(|\mathbf{r}_i-\mathbf{r}_j|\right) - \sum_{i,j} U^j_{en}\left(|\mathbf{r}_i-\mathbf{R}_j|\right) - \sum_{i,h,j} U^j_{een}\left(\mathbf{r}_i,\mathbf{r}_h,\mathbf{R}_j\right)\right\} \quad (2.2)$$

in which we include an electron-electron ee correlation term $U_{ee}\left(|\mathbf{r}_i-\mathbf{r}_j|\right)$, an election-nucleus en term $U^j_{en}\left(|\mathbf{r}_i-\mathbf{R}_j|\right)$ and a three body electron-electron-nucleus een term $U^j_{een}\left(\mathbf{r}_i,\mathbf{r}_h,\mathbf{R}_j\right)$. These functions are then expanded in fifth-order polynomials of the ee, en distances. The Jastrow factor is adapted to deal with pseudo-atoms. The Slater determinants, $D^{\uparrow}$ and $D^{\downarrow}$, are constructed from the sets of molecular orbitals $\{\phi^{\uparrow}\}$ and $\{\phi^{\downarrow}\}$ for the up- and down-spin electrons, respectively. We employ scalar-relativistic energy-consistent Hartree-Fock pseudopotentials specifically constructed for QMC calculations and expand the molecular orbitals on the corresponding cc-pVDZ basis set [68,68]. For the hydrogen atom, we use a more accurate BFD pseudopotential and basis set from M. Dolg and C. Filippi, unpublished. These pseudopotentials have been extensively benchmarked, and their reliability has been recently supported by a DMC computation of atomization energies to near-chemical accuracy [47]. The pseudopotentials are treated beyond the locality approximation [69].

Our choice of such minimal wavefunction and basis set is intentional since we want to maximize the scalability of our approach to systems larger than the ones considered here. Our interest here is not to challenge quantum chemistry methods for small systems but, rather, to devise a strategy that has a more extended range of applicability while preserving a notable accuracy.

2.3 Methodology: Wavefunction optimization

While most DMC calculations found in literature use molecular orbitals computed with some other electronic structure method [70–76], most often DFT, a key feature of our approach is that it is fully consistent since, at each iteration step in our geometric optimization, we perform a QMC optimization of all wave function parameters. This is done in order to guarantee consistency between the forces and the PES (see following sections) as well as to improve the results in terms of the absolute energy. It is found that this optimization procedure only approximately doubles the computer time needed to perform the calculations, while significantly lowering the expectation value of the energy. Full optimization of the wavefunction parameters is performed employing the strategy of energy minimization implemented in the CHAMP package

and described in [77]. The wavefunction in general has a non linear dependence on its parameters. However, we consider the expansion to linear order of the wavefunction

$$\begin{aligned} \psi_{lin}(\boldsymbol{\alpha}, \mathbf{r}) &= \psi(\boldsymbol{\alpha}^0, \mathbf{r}) + \sum_{i=1}^{N_\alpha} \Delta\alpha_i \psi_i'(\boldsymbol{\alpha}, \mathbf{r}) \\ \psi_i'(\boldsymbol{\alpha}, \mathbf{r}) &= \left. \frac{\partial \psi(\boldsymbol{\alpha}, \mathbf{r})}{\partial \alpha_i} \right|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}^0}, \end{aligned} \tag{2.3}$$

where $\boldsymbol{\alpha}^0$ are the initial parameters, $\Delta\boldsymbol{\alpha} = \boldsymbol{\alpha} - \boldsymbol{\alpha}^0$ are the parameters variation, N_α is the number of parameters. For linear parameters, the variations minimizing the energy are solution of the eigenvalue equation [78]

$$\boldsymbol{H}\Delta\boldsymbol{\alpha} = E\boldsymbol{S}\Delta\boldsymbol{\alpha}, \tag{2.4}$$

where $\boldsymbol{H}$ and $\boldsymbol{S}$ are the Hamiltonian and overlap matrix in the basis formed by the current wavefunction and its derivatives $\{\psi_0, \psi_1', \psi_2', \ldots, \psi_{N_\alpha}\}$. In a Monte Carlo sample these are estimated as:

$$\boldsymbol{H}_{ij} = \left\langle \frac{\psi_i'}{\psi_0} \frac{\mathcal{H}\psi_j}{\psi_0} \right\rangle_{\psi_0^2}; \quad \boldsymbol{S}_{ij} = \left\langle \frac{\psi_i'}{\psi_0} \frac{\psi_j}{\psi_0} \right\rangle_{\psi_0^2}, \tag{2.5}$$

where with the symbol $< \ldots >_{\psi_0^2} = \int \psi_0^2(\mathbf{r}) \ldots d\mathbf{r}$ we denote the average over the sampling of ψ_0^2. By solving Eq. 2.4 one can obtain linear variations $\Delta\boldsymbol{\alpha}$, however, as these are calculated in the first order approximation the parameters $\Delta\boldsymbol{\alpha} + \boldsymbol{\alpha}^0$ may be even worse than the original ones. In order to devise a better strategy, one possibility is to alter the dependence of the wavefunction on the nonlinear parameters by acting on the normalization of $\psi(\mathbf{r})$. A function $C(\boldsymbol{\alpha})$ is introduced, depending on the set of nonlinear parameters only, such that $C(\boldsymbol{\alpha}^0) = 1$, and a new wavefunction is defined $\tilde{\psi}(\boldsymbol{\alpha}, \mathbf{r}) = C(\boldsymbol{\alpha})\psi(\boldsymbol{\alpha}, \mathbf{r})$. The new derivatives are

$$\begin{aligned} \tilde{\psi}_i' &= \psi_i' + C_i'\psi_0 \\ C_i' &= \left. \frac{\partial C(\boldsymbol{\alpha})}{\partial \alpha_i} \right|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}^0}. \end{aligned} \tag{2.6}$$

We can then calculate the new variations $\Delta\tilde{\boldsymbol{\alpha}}$, and estimate the new $\tilde{\psi}_{lin} = \psi(\boldsymbol{\alpha}^0, \mathbf{r}) + \sum_{i=1}^{N_\alpha} \Delta\tilde{\alpha}_i \tilde{\psi}_i'(\boldsymbol{\alpha}, \mathbf{r})$. Since the variational space is the same and we are in the linear

approximation, $\tilde{\psi}_{lin}$ must be proportional to ψ_{lin}, i.e.:

$$\Delta\tilde{\boldsymbol{\alpha}} = \frac{\Delta\boldsymbol{\alpha}}{1 - \sum_{i=1}^{N_\alpha} C_i' \Delta\alpha_i}. \tag{2.7}$$

Imposing $C(\boldsymbol{\alpha})$ such that $\tilde{\psi}_i' \perp \psi_0$ one obtains good parameters but can lead to arbitrarily large parameters variations in some cases [79]. Another possible choice is to choose $C(\boldsymbol{\alpha})$ such that $\tilde{\psi}_i' \perp \psi_{lin}$ minimizes the linear wavefunction change $\left|\tilde{\psi_{lin}} - \psi_0\right|$, but this can lead to arbitrarily small parameters variations even far from the minimum. It is found that a reasonable choice is to impose an intermediate condition: $\tilde{\psi}_i' \perp \left(\frac{\psi_0}{2|\psi_0|} + \frac{\psi_{lin}}{2|\psi_{lin}|}\right)$. The parameter variations $\Delta\alpha_i$ are found to be unstable if the Monte Carlo sample is not large, In order to correct this problem, it is possible to add a positive constant a_{diag} to the Hamiltonian matrix $\boldsymbol{H}_{ij}$, except for the first element. As a_{diag} is made larger parameter variations are smaller and point toward the steepest descent direction. In the CHAMP code, once matrices **H** and **S** are computed, three different values of a_{diag} are used in computing parameter variations, differing by a factor of ten. The three resulting energies are computed by a correlated sampling run, and a parabolic fit determines (within certain bounds) the optimal value of a_{diag} used for the actual computations of the new parameter. After parameters are updated, the linear optimization procedure is iterated until convergence of the resulting energy. It may seem that the possibility of optimizing the wavefunction will be limited to small systems, as there is the need of storing matrices $\boldsymbol{H}$ and $\boldsymbol{S}$, $N_\alpha \times N_\alpha$ large. However, it was recently demonstrated that the optimization of large sets of parameters is possible [80], through the use of Krylov subspace algorithms, although we do not use this approach here.

2.4 Calculation of forces via correlated sampling

The computation of the reactant and product geometries, the NEB calculations, and the saddle-point location through the climbing-image method are all optimization procedures over the total energy, although with different constraints. They require, explicitly or implicitly, the ability to calculate derivatives of the total energy. We will be employing here forces calculated by finite increments via correlated sampling [52]. In principle it is possible and advisable to use forces by explicit differentiation in order to reduce the computational effort [54], however for the small systems reported here, the use of a finite-difference techniques is not coming with a large computational cost. We calculate forces employing a primary, reference configuration P and a secondary,

displaced perturbed configuration S. The force then will be:

$$F = \lim_{\Delta R \to 0} \frac{E^P - E^S}{\Delta R} \approx \frac{E^P - E^S}{\Delta R} \text{ for small } \Delta R, \tag{2.8}$$

where E is the total energy of the systems, and ΔR the ion coordinates displacement. The calculation of 2.8 can be easily computed numerically in a non stochastic approach by explicitly computing the energy of configurations P and S in separate runs, as long as the limit is not done literally and numerator and denominator of 2.8, last term, do not go below machine precision. In a stochastic approach however this is not possible: if we perform the calculation of E^P and E^S separately and evaluate the numerator of 2.8, we realize that we are subtracting two very similar quantities. Because these quantities have a statistical uncertainty σE^P and σE^S we see that $\sigma F \to \infty$ as $\Delta R \to 0$. However, a technique for dealing with such differences in a stochastic approach exists and is the so-called correlated sampling. This essentially consist in using the same sampling for systems S and P, and introducing in the averages a reweighting factor. This is done in order to exploit the cancellation of statistical errors arising from the use of the very same sampling. In VMC (see appendix B.1), one can then express the difference of energies as:

$$\begin{aligned} E^P - E^S &= \frac{1}{N} \sum_{i=1}^{N} -E_l^S \left(\psi^S \left(\mathbf{r}_i^P\right)\right) W_i + E_l^P \left(\psi^P \left(\mathbf{r}_i^P\right)\right) \\ W_i &= \frac{N \left|\psi^S \left(\mathbf{r}_i^P\right) / \psi^P \left(\mathbf{r}_i^P\right)\right|^2}{\sum_{j=1}^{N} \left|\psi^S \left(\mathbf{r}_j^P\right) / \psi^P \left(\mathbf{r}_j^P\right)\right|^2}, \end{aligned} \tag{2.9}$$

where E_l^P, E_l^S are the local energy terms of the two systems, Eq. B.2, and $\mathbf{r}_i^P$ denotes the electronic coordinates sampled over the wavefunction of the reference configuration $\psi^P(\mathbf{r}_i)$. W_i are the correcting weights we have to multiply the local energy terms of configuration S over the sampling of P, in order to obtain the correct result. Eq. 2.9 returns zero exactly (no statistical uncertainty) if $\mathcal{H}^P = \mathcal{H}^S$, e.g. the ionic configuration is the same, and $\psi^P(\mathbf{r}_i) = \psi^S(\mathbf{r}_i)$. Expression 2.9, inserted into Eq. 2.8, ensures that σF will not diverge anymore for $\Delta R \to 0$, but will actually go to a finite value. This expression however has to be further modified in order to obtain an efficient algorithm. The sampling of the electronic coordinates performed over system P will have a density peaked at ionic positions $\mathbf{R}_h^P$. This of course is not a proper distribution for the displaced ionic coordinates $\mathbf{R}_h^S$, introducing fluctuations in the W_i terms which eventually are reflected in a larger σF. This problem however can

be solved by introducing the so-called "space-warp" coordinate transformation [52], which drags the sampling points $\mathbf{r}_i$ for the configuration S so that, in the vicinity of a nucleus $\mathbf{R}_h^P$, they move rigidly with it as it is displaced to the new position $\mathbf{R}_h^S$.

$$\begin{aligned} \mathbf{r}_i^S &= \mathbf{r}_i^P + \sum_{h=1}^{N_{atoms}} \left(\mathbf{R}_h^S - \mathbf{R}_h^P\right) \omega_h(\mathbf{r}_i^P) \\ \omega_h(\mathbf{r}_i^P) &= \frac{F\left(\left|\mathbf{r}_i^P - \mathbf{R}_h^P\right|\right)}{\sum_{h=1}^{N_{atoms}} F\left(\left|\mathbf{r}_i^P - \mathbf{R}_h^P\right|\right)}, \end{aligned} \tag{2.10}$$

where $F(r)$ is a decaying function like r^{-k} or e^{-kr}, such that its decaying length is smaller of the typical interatomic distance. It is found that the choice of F has no major effect on performances, in the following r^{-4} will be used. With this modification, Eq.2.9 becomes:

$$\begin{aligned} E^P - E^S &= \frac{1}{N} \sum_{i=1}^{N} -E_l^S\left(\psi^S\left(\mathbf{r}_i^S\right)\right) W_i + E_l^P\left(\psi^P\left(\mathbf{r}_i^P\right)\right) \\ W_i &= \frac{N\left|\psi^S\left(\mathbf{r}_i^S\right)/\psi^P\left(\mathbf{r}_i^P\right)\right|^2 J\left(\mathbf{r}_i^P\right)}{\sum_{j=1}^{N}\left|\psi^S\left(\mathbf{r}_j^S\right)/\psi^P\left(\mathbf{r}_j^P\right)\right|^2 J\left(\mathbf{r}_j^P\right)}, \end{aligned} \tag{2.11}$$

where $J(\mathbf{r})$ is the Jacobian of transformation 2.10.

An additional complication arises due to the behavior of the local energy near the nodes. If we make the derivative of Eq. B.1 in the case where the operator is the Hamiltonian of the system we obtain

$$\begin{aligned} \langle E\rangle &= \frac{1}{Z}\int \mathrm{d}\mathbf{r}\ |\psi_T(\mathbf{r})|^2 E_l(\mathbf{r})\,; \quad Z = \int \mathrm{d}\mathbf{r}\ |\psi_T(\mathbf{r})|^2 \\ \left\langle \frac{dE}{dR}\right\rangle &= \frac{1}{Z}\int |\psi_T(\mathbf{r})|^2 \left(\frac{dE_l}{dR} + (E_l - \langle E\rangle)\frac{d\,ln(|\psi_T(\mathbf{r})|^2)}{dR}\right) \mathrm{d}\mathbf{r}. \end{aligned} \tag{2.12}$$

Near the nodes $E_l \propto \frac{1}{d}$, where d is the distance from the nearest nodal surface, $\frac{dE_l}{dR} \propto \frac{1}{d^2}$ and $\Pi \propto d^2$. As a result, the integral 2.12 is bounded. However, the variance of $\left\langle \frac{dE}{dR}\right\rangle$ is not, since

$$\begin{aligned} \sigma\left\langle \frac{dE}{dR}\right\rangle = \frac{1}{Z}\int |\psi_T(\mathbf{r})|^2 \left(\frac{dE_l}{dR} + (E_l - \langle E\rangle)\frac{d\,ln(|\psi_T(\mathbf{r})|^2)}{dR}\right)^2 \mathrm{d}\mathbf{r} - \\ \left(\frac{1}{Z}\int |\psi_T(\mathbf{r})|^2 \left(\frac{dE_l}{dR} + (E_l - \langle E\rangle)\frac{d\,ln(|\psi_T(\mathbf{r})|^2)}{dR}\right) \mathrm{d}\mathbf{r}\right)^2 \end{aligned} \tag{2.13}$$

and the first right hand side term $\left\langle \left(\frac{dE}{dR}\right)^2 \right\rangle$ is not bounded. This is also true in the finite difference case. In order to mend this problem, a solution was proposed in [56]. It is based on introducing an importance sampling to evaluate Eq. 2.12, with a the new sampling function defined as:

$$\begin{aligned} \psi_G(\mathbf{r},\epsilon) &= d(\mathbf{r})c(\mathbf{r},\epsilon)\psi_T(\mathbf{r}) \\ c(\mathbf{r},\epsilon) &= \begin{cases} \frac{1}{d(\mathbf{r})}, & \text{if } d(\mathbf{r}) \geq \epsilon \\ \epsilon \left(d(\mathbf{r})\epsilon\right)^{-\frac{1}{d(\mathbf{r})\epsilon}}, & \text{if } d(\mathbf{r}) < \epsilon \end{cases}, \end{aligned} \tag{2.14}$$

where $d(\mathbf{r})$ is a function which near the nodes reduces to the distance to the nearest node, otherwise has a finite value, and ϵ is a cutoff parameter. Notice that ψ_G reduces to ψ_T far from the nodes ($d(\mathbf{r}) \geq \epsilon$), while at the node ψ_G goes to a finite value. A suitable choice for $d(\mathbf{r})$ might be:

$$d(\mathbf{r}) = \left| \frac{\psi_T(\mathbf{r})}{|\bigtriangledown \psi_T(\mathbf{r})|} \right|. \tag{2.15}$$

Then, by dividing and multiplying into Eq. 2.12 by ψ_G, we obtain a new sampling expression for the local energy derivative:

$$\begin{aligned} Z' &= \int |\psi_G(\mathbf{r})|^2 \frac{|\psi_T(\mathbf{r})|^2}{|\psi_G(\mathbf{r})|^2} \, \mathrm{d}\mathbf{r} \\ \left\langle \frac{dE}{dR} \right\rangle &= \frac{1}{Z'} \int |\psi_G(\mathbf{r})|^2 \frac{|\psi_T(\mathbf{r})|^2}{|\psi_G(\mathbf{r})|^2} \left(\frac{dE_l}{dR} + (E_l - \langle E \rangle) \frac{d\, ln(|\psi_T(\mathbf{r})|^2)}{dR} \right) \mathrm{d}\mathbf{r}, \end{aligned} \tag{2.16}$$

which we can interpret stochastically by sampling over $|\psi_G(\mathbf{r})|^2$ and considering the rest of the integrand as the estimator. Since now our estimator is not diverging at the nodes, there is no problem of infinite variance anymore.

Also in DMC is possible to follow a similar route as in VMC, but complications are found. In DMC (see appendix B.2) we have to consider a primary walk generated for the reference configuration according to Eq. B.9, and a secondary walk for the displaced nuclei, generated from the primary one by applying the space warp transformation on each sampling point. As in VMC, the two configurations must have different wavefunctions ψ^P and ψ^S, as well as different Hamiltonians, resulting in different expression for DMC Green functions (Eq. B.9), which will be labeled by adding a superscript index P or S, like G^P and G^S. Therefore we have to correct for the wrong dynamics of the random walk of system S, as it is derived from the

one of system P. The secondary walk is generated according the drift-diffusion term (Eq. B.10) of the primary system $G^P_{dd}(\mathbf{r}^P, \mathbf{r'}^P, \tau)/J(\mathbf{r}^P)$, while a correct generation should have followed $G^S_{dd}(\mathbf{r}^S, \mathbf{r'}^S, \tau)$. In the presence of the Metropolis step Eq. B.12, the secondary move was accepted with probability P^P rather than P^S. Therefore, to correct for the wrong G_{dd}, the weight of the secondary walkers should have been multiplied by the term:

$$\frac{P^S\, G^S_{dd}(\mathbf{r}^S, \mathbf{r'}^S, \tau)}{P^P\, G^P_{dd}(\mathbf{r}^P, \mathbf{r'}^P, \tau)/J(\mathbf{r}^P)}. \tag{2.17}$$

It is found that this expression has strong fluctuations, making its application unpractical. Here we employ the approximate strategy described in [52], which essentially replaces expression 2.17 with the ratio of secondary and primary wave function $\left|\psi^S(\mathbf{r}^S)/\psi^P(\mathbf{r}^P)\right|^2 J(\mathbf{r}^P)$ as it is done in VMC, Eq 2.11. This is correct in the limit of a trial function equal to the true ground state wavefunction ψ_g, as in the case the growth/decay term does not modify walker's weight as it becomes simply $\delta(\mathbf{r}-\mathbf{r}')$ by an appropriate choice of E_l, and G_{dd} exactly samples ψ_T. In other words, in the presence of the acceptance step, the DMC procedures exactly reduces to a VMC procedure if $\psi_T = \psi_g$, with a proposal term given by G_{dd}, therefore in this limit resampling by Eq 2.11 is exactly equivalent to using 2.17. The approximation than lies in assuming that this substitution is reasonably good even if $\psi_T \neq \psi_g$. Of course, in this case the growth-decay term is not a simple delta and the weight of the secondary walkers must be propagated by according to $G^S_{g/d}(\mathbf{r}^S, \mathbf{r'}^S, \tau)$. In summary, the procedures follows the following steps:

i) A primary walk is generated by the standard DMC procedure.

ii) A secondary walk is generated from the primary by applying the space warp transformation on each element of the primary walk.

iii) Secondary weights at time step t ($w^S(t)$) are obtained from the primary ones ($w^P(t)$) by

$$w^S(t) = w^P(t) \sum_{i=1}^{N_{proj}} exp\left[-\tau\left(E_l\left(\mathbf{r}^S(t-i\tau), \psi^S\right) - E_l\left(\mathbf{r}^P(t-i\tau), \psi^P\right)\right)\right] \tag{2.18}$$

N_{proj} is a number of time steps large enough to project out the secondary ground state, but small enough to avoid large fluctuations of $w^S(t)$.

iv) The difference in energy is calculated as in 2.11, the simple average being replaced

by a weighted average according to weights calculated in the previous point:

$$E^P - E^S = -\frac{\sum_{t=t_0}^{t_0+N} w^S(t) E_l^S\left(\psi^S\left(\mathbf{r}^S(t)\right)\right) W(t)}{\sum_{t=t_0}^{t_0+N} w^S(t)} + \frac{\sum_{t=t_0}^{t_0+N} w^P(t) E_l^P\left(\psi^P\left(\mathbf{r}^P(t)\right)\right)}{\sum_{t=t_0}^{t_0+N} w^P(t)}$$
$$W(t) = \frac{N\left|\psi^S\left(\mathbf{r}^S(t)\right)/\psi^P\left(\mathbf{r}^P(t)\right)\right|^2 J\left(\mathbf{r}^P(t)\right)}{\sum_{t'=t_0}^{t_0+N}\left|\psi^S\left(\mathbf{r}^S(t')\right)/\psi^P\left(\mathbf{r}^P(t')\right)\right|^2 J\left(\mathbf{r}^P(t')\right)}. \tag{2.19}$$

It is found that this approximation is surprisingly accurate [52]. Also in DMC we have to use the guiding function approach already shown for the VMC case in our sampling in order to avoid infinite variance in the energy derivatives estimators, Eq. 2.14. By employing a time step dependent cutoff $\epsilon(\tau)$ such that $\epsilon(0) = 0$, the procedure is correct in the limit of $\tau \to 0$.

When calculating forces by using finite differences, Eq. 2.8, we can improve the estimate by calculating the force both for ΔR and $-\Delta R$, and then averaging over the result $\frac{F(-\Delta R)+F(\Delta R)}{2}$. This symmetrization of the estimate ensures that the procedure is correct to $O(\Delta R^2)$ instead of $O(\Delta R)$, and allows the estimation of the second derivative of the energy along the direction defined by ΔR as $\frac{F(-\Delta R)-F(\Delta R)}{\Delta R}$, which is a useful quantity for minimization procedures we intend to use.

2.5 Force bias from wavefunction choice

We are calculating forces by evaluating the difference in energy of two configurations. This implies that we have to write two different wavefunctions $\psi^P(\mathbf{r})$ and $\psi^S(\mathbf{r})$, as the energy evaluated in VMC/DMC depends through ionic coordinates and wavefunction: $E^S\left(\psi^P(\mathbf{r}), \mathbf{R}^P\right)$. In principle, in the Born-Oppenheimer approximation, we would like E^P and E^S to be the ground state energies; however this would imply that we knew the exact ground state wavefunction. In reality, we are restricting the form of the wavefunction to be a linear combination of Slater determinant of molecular orbitals expressed in some basis multiplied by a Jastrow factor. The VMC/DMC energy of a configuration depends on the set of wavefunction parameters (basis and Jastrow coefficients), that we will collectively indicate with the Greek letter α. The optimal VMC/DMC parameter in turn depend on the ionic coordinates $\mathbf{R}$, therefore we can write $E(\alpha_i\{\mathbf{R}\}, \mathbf{R})$. Of course the choice of the wavefunction influences the accuracy on forces. Here we deal with an additional source of error, which depends on the way we generate the wavefunction for the displaced configuration $\psi^S(\mathbf{r})$. As ionic displacements are by definition small, a common choice to generate the secondary

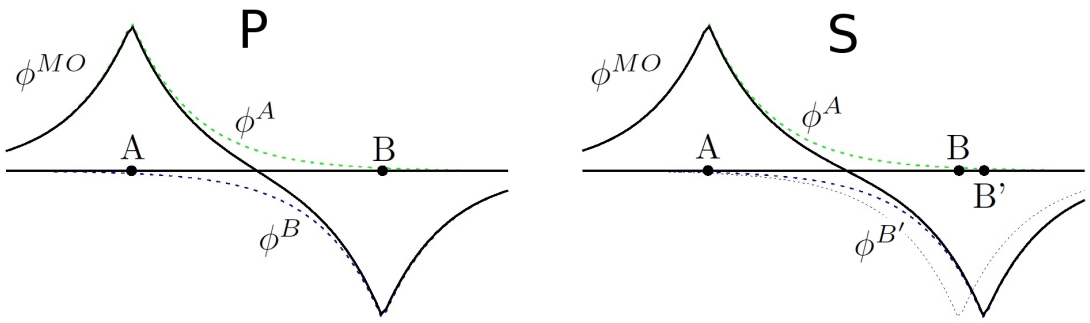

Figure 2.1: Schematic representation of the procedure for generating the perturbed wavefunction needed for correlated sampling. In configuration P the molecular orbital ϕ^{MO} is generated combining linearly atomic orbitals ϕ^A and ϕ^B. In configuration S atom B is moved to B'. The atomic orbital of atom B ϕ^B is displaced rigidly with the nucleus in position B' without being modified and a new molecular orbital is generated accordingly. Jastrow factor is kept fixed.

wavefunction is to take the wavefunction for system P and recenter the atomic orbitals into the new nuclei position without changing the coefficients of their expansion in the basis set and keeping the Jastrow fixed. This, however, introduces an error in the calculation of correlated sampling VMC/DMC forces, as we would like to calculate the total derivative of the energy with respect to the ionic displacements, compatibly with the form of the employed wavefunction. Suppose we want to calculate the force along some direction R_h, this total derivative would be:

$$\frac{dE\left(\alpha\{\mathbf{R}\},\mathbf{R}\right)}{dR_h} = \frac{\partial E\left(\alpha\{\mathbf{R}\},\mathbf{R}\right)}{\partial R_h} + \sum_j \frac{\partial E\left(\alpha\{\mathbf{R}\},\mathbf{R}\right)}{\partial \alpha_j}\frac{d\alpha_j}{dR_h}. \tag{2.20}$$

If we neglect the variation of parameters with ionic displacements, we are effectively neglecting the second right hand side term of equation 2.20, therefore we are taking only a partial derivative of the energy instead of the total one. One solution is to optimize the wavefunction VMC/DMC parameters in configuration P and S and use these wavefunctions in Eq. 2.11. However Eq. 2.20 tells us that at first order we need only to optimize $\psi^P\left(\mathbf{r}\right)$ to get an unbiased result, as by definition for optimized parameters

$$\frac{\partial E\left(\alpha\{\mathbf{R}\},\mathbf{R}\right)}{\partial \alpha_j} = 0 \;\; \forall j. \tag{2.21}$$

In this case the second right hand side term of Eq. 2.20 vanishes and $\psi^S(\mathbf{r})$ can be generated by recentering the basis functions and neglecting parameter variations.

In order to estimate the bias incurred neglecting the last term of Eq. 2.20 when the wavefunction parameters are not optimal, we tested the consistency between the force obtained with the correlated sampling method and the one obtained fitting the potential energy curve for a C_2 molecule, for different types of wavefunctions and levels of optimization. Optimization of the wavefunction parameters is done employing the procedure described in section 2.3. The C_2 molecule was chosen because it is a relatively cheap system to perform calculations on, yet is not trivial for its multideterminantal character, therefore is representative of a challenging system for QMC calculations. Let's start from a non optimized case. We take the C_2 molecule and calculate, using the GAMESS package (with the same pseudopotentials and basis set described in section 2.2), the HF wavefunction at the experimental equilibrium distance of the molecule and, using this determinant, optimize a Jastrow factor at VMC level via the CHAMP code. We then calculate at various interatomic distances the HF orbitals, and perform a VMC/DMC calculation of energy and force using the corresponding HF determinant coupled to the same Jastrow factor obtained as described before. The comparison between the force obtained by fitting the potential energy curve with a third order polynomial (and computing the derivative) with the force obtained with correlated sampling is reported in Figure 2.2, panel a). We observe that in this case there is no consistency, with the force from the correlated sampling being far from the one obtained by fitting the PES, due to the poorness of the wavefunction. If we instead of using an HF calculation we use a better wavefunction, such as the one obtained by replacing the HF calculation with a B3LYP DFT calculation in the aforementioned procedure, we see that the agreement improves, Fig.2.2, panel b). However, there is still no consistency, as the data do not overlap even considering their statistical uncertainty. In Fig.2.2, panel c), we optimize the wavefunction (molecular orbitals and Jastrow) at each considered distance at VMC level. In this case there is a full consistency between the line and the datapoint, as expected. It is interesting to check whether using more sophisticated wavefunctions than single determinant ones is possible to obtain a reasonable agreement between correlated sampling force and PES without the need of a VMC optimization. In this case, we employ a simple complete active space (CAS) wave function correlating three electrons in the three more relevant orbitals, calculated self-consistently with the GAMESS package at each distance tested. In case Fig.2.2, panel d), we couple this wavefunction with a fixed (i.e. not changing with the various interatomic distances

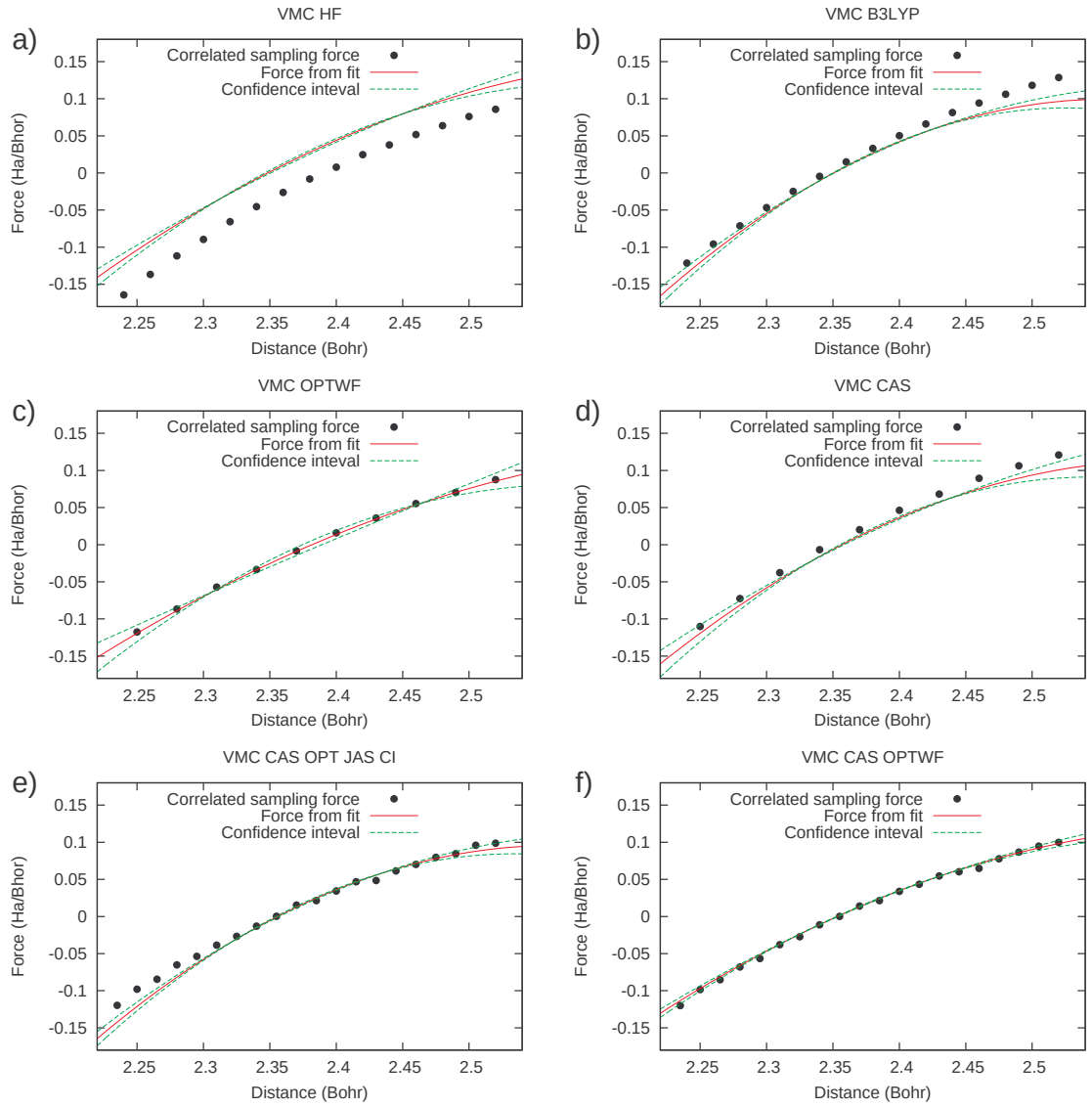

Figure 2.2: Interatomic VMC force for the C_2 molecule as obtained with various methods and levels of wavefunction optimization. Force obtained with correlated sampling (datapoints, statistical error within circle size) is compared with the force obtained by fitting the VMC potential energy curve and computing the derivative (full line with dashed confidence interval). Panels correspond to the following wavefunctions:
a) HF determinant calculated at each distance, fixed Jastrow optimized at the experimental equilibrium distance.
b) B3LYP determinant calculated at each distance, fixed Jastrow optimized at the experimental equilibrium distance.
c) determinant and Jastrow optimized at each distance.
d) CAS wavefunction calculated at each distance, fixed Jastrow optimized at the experimental equilibrium distance.
e) CAS wavefunction calculated at each distance, optimized Jastrow and CI expansion coefficients at each distance.
f) CAS wavefunction fully optimized (orbitals, CI coefficients, Jastrow) at each distance.

tested) Jastrow, optimized at the equilibrium distance of the molecule as usual, and use it for a VMC energy/force calculation. We see that, although the wavefunction is of better quality and retrieves a lower energy than the single determinant optimized ones (-11.0427 ± 0.0004 vs. 11.0240 ± 0.0004 Ha at the equilibrium distance), it does not return a consistent force. This is because the condition 2.21 has not any direct relation to the overall quality of the wavefunction, but only with the optimization of its parameters. Even a partial optimization of the CAS wavefunction is not enough to retrieve a full consistency. This is shown in Fig.2.2, panel e), were we repeat the procedure of case d) but we precede the energy/force calculation with an optimization of the determinantal expansion (CI coefficients) and Jastrow parameters at VMC level at each distance, keeping thus fixed the molecular orbitals in the determinants. In Fig.2.2, panel f), full optimization of the CAS wavefunction (molecular orbitals, CI coefficients, Jastrow) is performed at each distance, thus recovering the expected agreement. We attribute the small scattering of data present in this case panel with a less than perfect wavefunction optimization.

We also repeated the same analysis substituting the final VMC energy/force calculation with a DMC one, but keeping the various wavefunction optimization at VMC level, as doing this in DMC would be too expensive in an actual case. By doing this procedure we expect two sources of error to be present. The first one is arising from the fact that we are enforcing condition 2.21 at VMC level, but we do not expect this condition to hold exactly at DMC level. The other source of error is due to the approximate reweighting procedure for the correlated sampling calculation that we are employing, described in section 2.4. Results are illustrated in Fig. 2.3, where the panels are completely analogous to the one of Fig. 2.2. We use a time step of 0.01 a.u. in the DMC calculations. The outcome appears to be pretty similar: if complete optimization of the wavefunction is not performed, there is no complete consistency between correlated sampling force and PES, irrespective of the number of determinants included into the wavefunction. In the fully optimized cases c) and f) we retrieve good results, even if the consistency appears to be less striking than in the VMC case: this is due to the sources of error pointed out before.

In conclusion, we have shown in this section that full wavefunction optimization is needed to obtain full consistency between forces calculated by correlated sampling in VMC/DMC and the potential energy curve, even for wavefunctions of good quality. However, the errors displayed here may be small enough for many application, therefore full wavefunction optimization may be not necessary. Additional errors are present in DMC calculations shown here, due to the approximate reweighting tech-

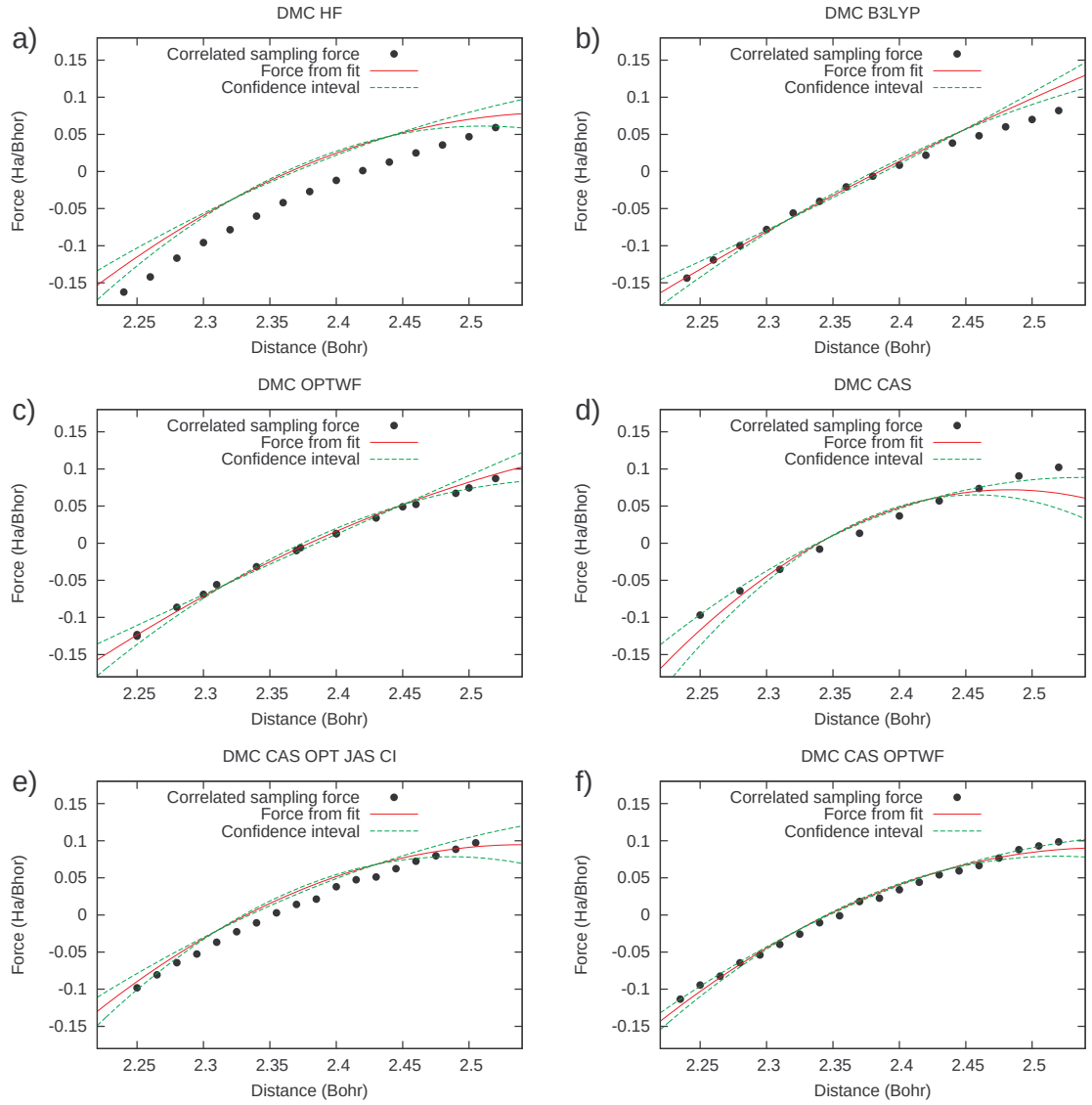

Figure 2.3: Interatomic DMC force for the C_2 molecule as obtained with various methods and level of wavefunction optimization. All optimizations are done at VMC level, *not at DMC level*. Force obtained with correlated sampling corresponds to datapoints, and is compared with the force obtained by fitting the DMC potential energy curve and computing the derivative corresponds to the full line with dashed confidence interval. Panels refer to the same wavefunctions as in Figure 2.2

.

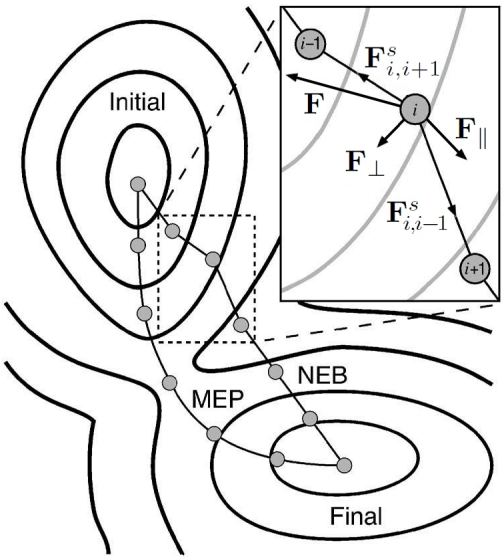

Figure 2.4: Illustrative example of a Minimum Energy Pathway (MEP) connecting two stable states in a Potential Energy Surface (PES) and of the forces driving the minimization of an initial guess of the Nudged Elastic Band (NEB) to the MEP (magnified panel).

nique employed and the fact that we did not perform optimization of the wavefunction ad DMC level but at VMC level, since optimization at DMC level would be exceedingly expensive in an actual application. However, the extent of these errors is small enough for most applications.

2.6 Nudged Elastic Band and Climbing Image methods

Reactants and product states of a chemical reaction are by definition stable: therefore, in configuration space, they sit in minima of the potential energy surface (PES). From this follows that the path of minimum energy (MEP) connecting these minima will go through one or more saddle points, the so-called transitions states. The saddle point are stationary, but are not stable as at least one eigenvalue of the Hessian matrix is negative. The Nudged Elastic Band method (NEB) is a local search method for the MEP connecting reactants (initial state) and products (final state) of a reaction [61]. It is a local search procedure in that it does not guarantee to converge to the true MEP (i.e. the one having the lowest saddle point), but in the space of configuration of the NEB recovers the energy pathway having a basin of attraction into which the

initial guess falls. Thus the initial guess mus be "near" enough to the true MEP in order to converge to it. Therefore the method is suitable for simple reactions, in which there is little doubt about the mechanism, or reactions in which the MEP is approximately already known. The idea behind the method is to discretize an initial guess for a MEP connecting the stable states into a series of "images", the first one being the initial state and the last one being the final, see Fig. 2.4. The first and last images are kept fixed, all the others evolve their position following a dynamics that let the NEB "fall" into the saddle point and valleys of the PES, while keeping the images roughly equispaced in configuration space. The commonly employed dynamics is driven by the following component of the force applied on each image i:

i) A force parallel to the NEB, $\mathbf{F}_{\parallel}$. The parallel direction $\mathbf{n}_{\parallel}$ is obtained by the normalized difference the versors pointing from image i to $i+1$ and $i-1$. A spring force, denoted with $\mathbf{F}^s_{i,i+1}$, is applied pointing toward image $i+1$, of modulus $K\left|R_i - R_{i+1}\right|^2$, and another is applied toward image $i-1$. The acting spring force is then projected on the parallel direction $\mathbf{F}_{\parallel} = \left(\left(\mathbf{F}^s_{i,i+1} + \mathbf{F}^s_{i,i-1}\right)\cdot \mathbf{n}_{\parallel}\right)\mathbf{n}_{\parallel}$. The effect of this force in the minimization procedure is simply to slide the images along direction $\mathbf{n}_{\parallel}$ until they are equispaced, independently of the spring constant K.

ii) A force perpendicular to the NEB, $\mathbf{F}_{\perp}$. We consider the force $\mathbf{F}$ applied by the PES on image i (minus the gradient of the PES), and remove from it the component parallel to the NEB, defined in **i)**, to obtain $\mathbf{F}_{\perp} = \mathbf{F} - \left(\mathbf{F}\cdot\mathbf{n}_{\parallel}\right)\mathbf{n}_{\parallel}$. This component is the one letting the NEB move toward the MEP.

These forces are represented in Fig. 2.4, magnified panel. Following these forces one applies a minimization procedure iteratively, until the NEB reaches equilibrium. In our case, we employ the Newton minimization method on the single images until convergence is reached.

In the equilibrium configuration, the NEB is an approximation of the MEP. There is no reason to believe that one of the images is lying exactly on the transition state. For this reason, another procedure, the so-called climbing image method, is commonly used to locate the saddle point. The procedure involves only the image with the highest energy i, plausibly the nearest to the saddle point, and its neighbouring images. The images $i+1$ and $i-1$ are kept fixed, and image i is evolved according to the forces components:

i) Perpendicular force $\mathbf{F}_{\perp} = \mathbf{F} - \left(\mathbf{F}\cdot\mathbf{n}_{\parallel}\right)\mathbf{n}_{\parallel}$.

ii) Parallel force $\mathbf{F}_{\parallel} = -\left(\mathbf{F}\cdot\mathbf{n}_{\parallel}\right)\mathbf{n}_{\parallel}$.

This simple procedure transforms the saddle point into a minimum for image i, as the force in the parallel direction, the one along which the curvature of the PES is negative, is reversed. Therefore, if the images $i+1$ and $i-1$ are near enough to the saddle point so that the direction connecting the two images is a sufficiently good description of the direction along which the curvature is negative, the image i converges to the transition state exactly. The NEB-climbing image procedure therefore provides with the geometry and energy of the transition state. This information, together with the energy of the initial/final states, allows calculation of the forward/reverse barrier of the reaction.

2.7 Results

We select four challenging reactions from the NHTBH38/04 database [44] plus one hydrogen transfer reaction. As best estimates, we use the atomic geometries for the initial, final, and transition states reported in the database and computed through a quadratic configuration interaction with single and double excitations (QCISD) optimization. For these geometries, the barrier heights estimated with the W1 method (a complete basis set extrapolation over CCSD(T)) are also available. The reference data for $H + OH \rightarrow H_2 + O$ are from ext-CAS+1+2+Q calculations [65].

We initially focus on the $H + F_2 \rightarrow HF + F$ reaction, in which the atoms are aligned through all the reaction. The DFT and Hartree-Fock (HF) all-electron calculations are performed with the GAMESS package [63], using Dunning-type Correlation Consistent triple-zeta basis sets, augmented with a set of diffuse function (aug-cc-pVTZ). No relativistic correction is included. QMC calculations include them through the scalar-relativistic pseudopotentials, however these corrections are small for the light elements considered here and will not affect our results. An example of the obtained NEB for this reaction is reported in Figure 2.5 for VMC and for the functional B3LYP, in the space of the H-F and F-F bond distances. In the figure it is possible to appreciate how the two methods return similar geometries near the equilibrium geometries, while differ significantly in the region where the actual reaction takes place. In particular, the VMC NEB passes much nearer to the best estimate saddle point than B3LYP one. We now focus on the three most significant geometries along the reaction: the initial, final and transition state ones; DFT and QMC data are collected in Table 2.1. To measure how much the geometries differ from the best estimates, we calculate the RMS deviations of the interatomic distance among all

atoms with respect to the corresponding best-estimate geometries. In the Table, a forward barrier (V_f) of zero means that the DFT functional finds no transition state (i.e. the reactants are unstable) with the reverse barrier being the reactant-product energy difference. Many DFT functionals fail in finding any transition state for this reaction, including the hybrid functionals PBE0 [81] and B3LYP, while M06 [82] retrieves a saddle point but with large deviations over the best estimate transition state geometry. In Table 2.1, we also report the initial/final/transition state geometries computed via VMC forces, where the uncertainty on the interatomic bonds due to the statistical noise on the forces is about 0.002 Å. These geometries come from a fully VMC NEB and climbing image calculations. Since forces calculated in QMC possess a statistical uncertainty, strictly speaking the optimization procedure via the Newton method never converges. Therefore, the equilibrium positions of the images along the nudged elastic band are obtained by averaging over several iterations after all quantities vary only by statistical fluctuations around a stationary value. VMC is able to retrieve even at the single-determinant level the initial, the final and, especially, the transition state geometry with much better accuracy than DFT. It is interesting to notice that, for geometrical data, VMC also performs better than the hybrid-meta M06 functional which is constructed to fit the barriers calculated on the best-estimate geometries from the database [82]. Clearly, this fitting procedure does not always guarantee that the actual transition state retrieved by the functional is near the best-estimate one. For testing purposes we also have located the saddle point using non optimized orbitals from the M06 functional and a fixed Jastrow; we find that the result is substantially worse, with a RMS deviation from the best estimate of 0.149 Å. This result confirms that at least in some cases the wavefunction optimization is needed to obtain accurate values.

Although the VMC geometries are rather accurate, the predicted reverse energy barrier (V_r) is markedly overestimated. Performing a DMC (time step of 0.01 a.u.) calculation on the VMC geometries retrieves better energy estimates, but the error on the reverse barrier of about 8 Kcal/mol is still quite significant. We find that, to improve this energy barrier, it is not useful to reoptimize the geometries via DMC forces: the use of DMC forces does not alter significantly the geometries (within a statistical uncertainty on the transition state geometry of about 0.008 Å) and, consequently, the estimated barrier heights either. In order to improve over these values, it is possible to take advantage of the variational principle available in QMC and to resort to the use of multi-determinantal wave function. In this case, we employ a simple complete active space (CAS) wave function correlating three electrons in

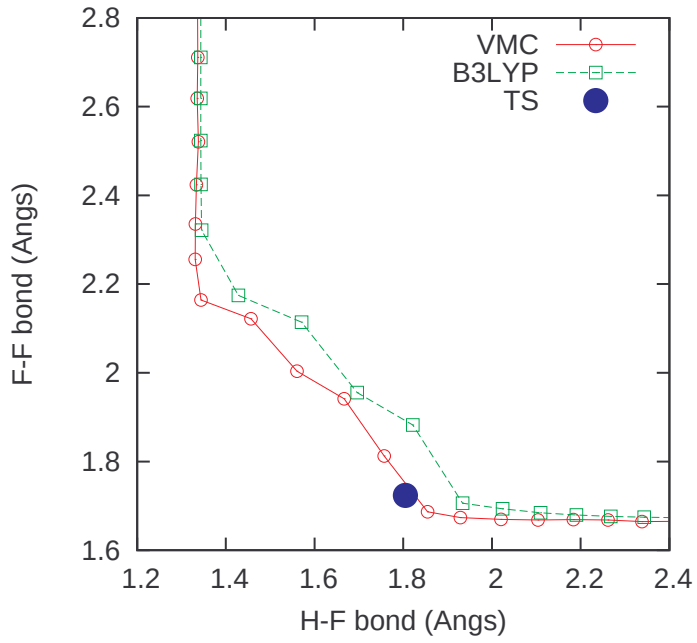

Figure 2.5: Nudged Elastic Band results obtained with single-determinant VMC (red full line) and with the functional B3LYP (green dashed line), represented in the space of the H-F and F-F bond distances (in this case the atoms are always aligned during the reaction). The best estimate transition state position is also reported (full circle).

the three active orbitals relevant for the reaction and recompute the energy barriers over the VMC geometries obtained at the single-determinant level. The resulting VMC barriers are improved and the DMC values become very similar to the best estimates. Although the use of CAS wave functions is not readily applicable to larger systems due to their exponential scaling with system size, there exist scalable techniques to improve over single-determinant wave functions through the design of accurate multi-determinantal size-extensive and linear-scaling [83] or backflow wave functions [45, 84, 85].

We now return to the simple Slater-Jastrow wave function, and consider the other reactions. In Figure 2.6, we compare the difference between the geometries obtained by VMC and various DFT functionals with respect to the reference ones obtained with QCISD, using the same measuring criterion as in Table 2.1. The functionals reported in this figure are all hybrid or meta-hybrid and are the ones returning the smallest geometric/energy deviation among the ones we tested on this set of reactions, which are the same ones listed in Table 2.1. The generalized-gradient-approximation

Table 2.1: Barrier heights and RMS of geometric deviations, $H + F_2 \rightarrow HF + F$ reaction, for reactants (React), products (Prod) and transition states (TS). We denote with V_f/V_r the forward/reverse reaction barrier heights (BHs). The RMS is calculated over the deviation of the interatomic distances of all the atoms from the best estimate geometry.

$H + F_2 \rightarrow HF + F$		BE	VMC	VMC CAS	DMC	DMC CAS	HF	LSDA	BLYP	B3LYP	PBE	PBE0	M06
BHs	V_f	2.27	6 ± 1	1.3 ± 0.2	2.2 ± 0.5	1.4 ± 0.3	0.0	0.0	0.0	0.0	0.0	0.0	2.7
(Kcal/mol)	V_r	106.18	126 ± 1	112.2 ± 0.6	114 ± 1	105.4 ± 0.7	127.3	83.2	90.9	100.9	87.9	100.0	107.8
RMS deviation (Å)	React		0.008		0.007		0.067	0.010	0.037	0.002	0.018	0.019	0.020
	Prod		0.002		0.008		0.016	0.018	0.020	0.009	0.017	0.004	0.001
	TS		0.028		0.013		-	-	-	-	-	-	0.216

functionals used in the overwhelming majority of DFT applications of CI-NEB, entail significantly larger deviations. For example, the deviation of PBE is two to four times larger than that of PBE0 for transition state geometries reported here, and the RMS barrier heights deviation of PBE on this set is more than twice than that of PBE0. For equilibrium geometries, VMC performs at the level of the hybrid functionals. For the transition state, it typically returns more accurate geometries, often performing much better than DFTs. Notwithstanding that the M06 is actually fitted to reproduce barrier heights for the NHTBH38/04 reactions, VMC still performs better than this functional in evaluating transition state geometries. This may be again due to the M06 parametrization procedure, which does not guarantee the accuracy of the actual transition state calculated by the functional. We do not recalculate the geometries employing DMC forces because, from the test performed, DMC forces improves VMC transition state geometries only slightly, as these are already notably accurate. Furthermore, DMC geometry corrections are barely reflected in the calculation of the barrier heights, as these energies are second order in the deviation from the actual equilibrium points. For these reasons, we speculate that the calculations of geometries at VMC level and of barrier heights at DMC level is the most sensible choice regarding the trade-off between accuracy and computational cost. In Figure 2.7, we report the forward and reverse reaction barriers. While VMC (not shown) performs less accurately than hybrid functionals, DMC calculated on the VMC geometries significantly improves the barrier heights estimates of all reactions upon the VMC values, and performs at the level of the hybrid DFT approaches. In particular, our QMC procedure is more accurate than the hybrid functionals B3LYP and PBE0, while the only functional performing on average at the same level or slightly better is M06, despite these barriers being calculated on transition state geometries worse than VMC ones. Moreover M06 is actually fitted to reproduce precisely the NHTBH38/04 bar-

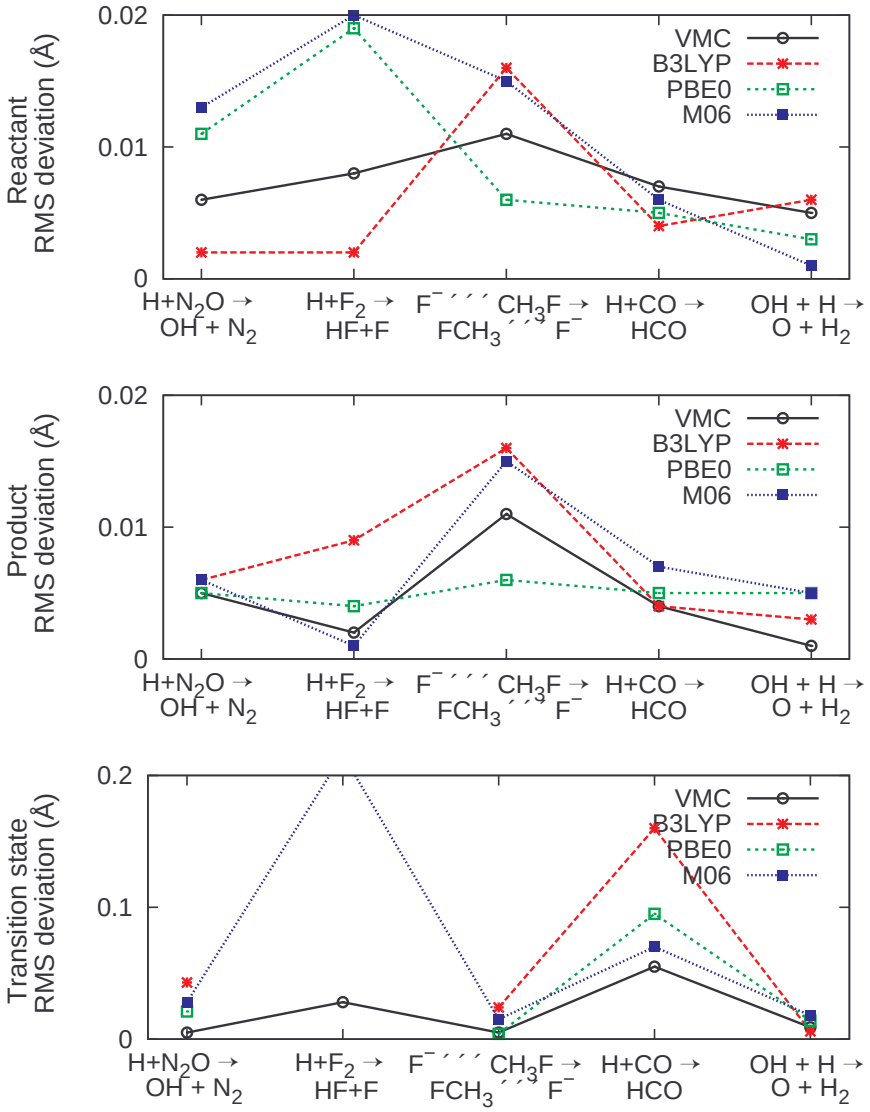

Figure 2.6: RMS of deviation of interatomic distances from the QCISD geometries (Å). Distances are calculated among all atoms involved in the reactions.

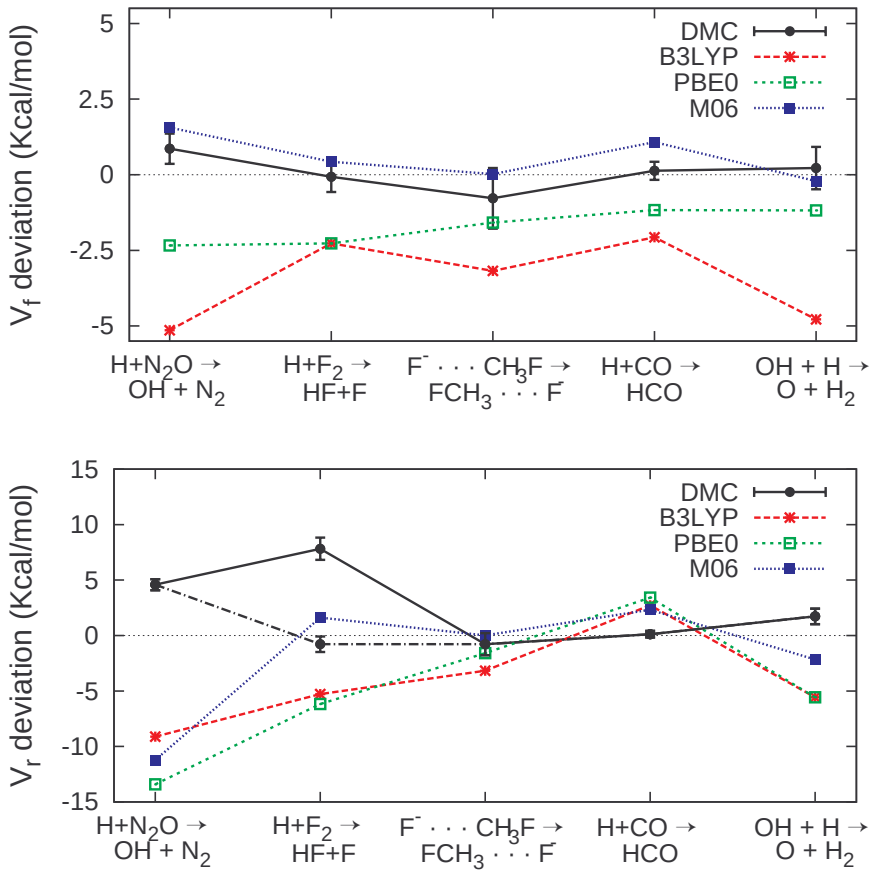

Figure 2.7: Forward (V_f) and reverse (V_r) barrier heights deviation from best estimates calculated by QMC and DFT Methods (Kcal/mol). The dash-dotted line represents the DMC values obtained by employing the CAS wavefunction for the $H + F_2 \rightarrow HF + F$ instead of the single determinant one (full line), see text.

rier heights, and there is no guarantee that, on a different set of reactions, it would still perform as well as QMC. Note that single-determinant DMC performs better than all DFT approaches, included M06, even in the reaction $H + N_2O \rightarrow OH + N_2$, which is known to be strongly multi-determinantal in character. Overall, QMC gets both the geometry and the energetics accurately, offering a parameter-free, more balanced description of reactants, products and transition states than all DFT schemes considered here.

2.8 Conclusions

We investigated the possibility of performing full-QMC reconstruction of minimum energy pathways of chemical reactions. Full optimization of the wave function parameters is carried out during each iteration, so the employed technique is internally fully consistent. Geometric optimization of the minimum energy pathway and of the transition state is done at VMC level, with the obtained geometries being more accurate than DFT ones, especially for transition states. It also demonstrates the ability to correctly locate the transition state in cases in which DFT fails in returning accurate geometries. At DMC level, the method displays very good performance in evaluating barrier reaction heights, comparing favorably even against hybrid functionals. Therefore, our approach of calculating the geometries at the VMC level and reaction barriers at the DMC level is most effective as far as performance over computational cost is concerned: calculating DMC geometries is very expensive and, in the tested cases, it does not improve significantly the estimates, while calculating DMC energies over VMC geometries is much cheaper and still retrieves good results. Since the employed wave function is of the simple Slater-Jastrow type, this technique is scalable to larger systems. Our results indicate that, for intermediate-sized system reactions where quantum chemistry methods are not computationally viable, the QMC approach may be the most accurate technique currently available.

Appendix A

A.1 Path Integral Monte Carlo

One of the leading techniques in the study of quantum many-body systems is the Path Integral Monte Carlo approach (PIMC). This ab-initio technique is capable of providing with unbiased estimates of physical observables at finite temperature on Bosonic systems. Being a Monte Carlo technique, the word *exact* in this context has to be intended in a numerical sense, meaning that all the errors (statistical and systematic) can in principle be driven to zero. Although the technique is very successful in the study of superfluidity in Bosonic systems, its applicability is strongly limited by the poor scaling with system size, as we will illustrate below.

PIMC originates from Feynman mapping of a quantum system onto a classical models of interacting "polymers" (path integral representation). This was subsequently translated in a Monte Carlo technique [86] based on sampling the diagonal part of the density matrix of a Bosonic system. On the continuum it was first applied to a system of 64 ^{4}He atoms in periodic boundary condition, in the pioneering work of Ceperley and Pollock [87]. We start by recalling that in thermal equilibrium properties of a quantum system can be calculated as:

$$\langle O \rangle = \frac{\sum_i \langle \phi_i \mid O \mid \phi_i \rangle \, e^{-\beta E_i}}{Z} = \frac{\sum_i \left\langle \phi_i \mid e^{-\beta H} O \mid \phi_i \right\rangle}{Z}, \tag{A.1}$$

where ϕ_i and E_i are the eigenvectors and eigenvalues of the Hamiltonian. $\beta = (K_b T)^{-1}$ is called the imaginary time or the inverse temperature. Z is the partition function:

$$Z = \sum_i e^{-\beta E_i}. \tag{A.2}$$

The thermal density matrix operator (we will always be dealing with its unnormalized

version) is $e^{-\beta H}$, or in eigenvector representation:

$$\rho\left(\phi_i, \phi_j, \beta\right) = \left\langle \phi_i \mid e^{-\beta H} \mid \phi_j \right\rangle. \tag{A.3}$$

Therefore Eq. A.1 can be written as:

$$\langle O \rangle = \frac{Tr\left(\rho O\right)}{Tr(\rho)}, \tag{A.4}$$

i.e. all the properties of a quantum system in equilibrium can be calculated from the thermal density matrix. This of course is true also in position representation. If we denote with $R = (r_1, r_2, ..., r_N)$ a vector containing the dN particles coordinates, d being the dimensionality of the system, we can write the density matrix in position representation as:

$$\rho\left(R', R'', \beta\right) = \left\langle R' \mid e^{-\beta H} \mid R'' \right\rangle \tag{A.5}$$

and Eq. A.1 becomes

$$\begin{gathered} \langle O \rangle = \frac{\sum_i \left\langle \phi_i \mid e^{-\beta H} O \mid \phi_i \right\rangle}{\sum_i \left\langle \phi_i \mid e^{-\beta H} \mid \phi_i \right\rangle} = \\ \frac{\int dR' dR'' dR''' \sum_i \left\langle \phi_i \mid R' \right\rangle \left\langle R' \mid e^{-\beta H} \mid R'' \right\rangle \left\langle R'' \mid O \mid R''' \right\rangle \left\langle R''' \mid \phi_i \right\rangle}{\int dR' dR'' \sum_i \left\langle \phi_i \mid R' \right\rangle \left\langle R' \mid e^{-\beta H} \mid R'' \right\rangle \left\langle R'' \mid \phi_i \right\rangle} = \\ \frac{\int dR' dR'' \left\langle R' \mid e^{-\beta H} \mid R'' \right\rangle \left\langle R'' \mid O \mid R' \right\rangle}{\int dR' \left\langle R' \mid e^{-\beta H} \mid R' \right\rangle} = \frac{\int dR' dR'' \rho\left(R', R'', \beta\right) \left\langle R'' \mid O \mid R' \right\rangle}{\int dR' \rho\left(R', R', \beta\right)}. \end{gathered} \tag{A.6}$$

Of course we do not know in general how to evaluate $\rho\left(R', R'', \beta\right)$. However, in many cases it is possible to express a sufficiently good approximation of the density matrix for small β. Therefore, one can make use of the property of the exponential operator $e^{-(\beta_1+\beta_2)H} = e^{-\beta_1 H} + e^{-\beta_2 H}$, which in position representation reads:

$$\rho\left(R', R'', \beta_1 + \beta_2\right) = \int dR''' \rho\left(R', R''', \beta_1\right) \rho\left(R''', R'', \beta_2\right). \tag{A.7}$$

Employing Eq. A.7 one can discretize the full imaginary time in M "slices" of length $\tau = \beta/M$ and represent the density matrix as a convolution in space of M density matrices.

$$\rho\left(R_0, R_M, \beta\right) = \int dR_1 dR_2 ... dR_M \rho\left(R_0, R_1, \tau\right) \rho\left(R_1, R_2, \tau\right) ... \rho\left(R_{M-1}, R_M, \tau\right). \tag{A.8}$$

The subscript index of the vector R will be referred to in the following as the time slice index. Now, in order to be able to express a short time approximation for $\rho(R_0, R_1, \tau)$ it is useful to split the kinetic and potential part of an Hamiltonian $\mathcal{H} = T + V$, in the so-called primitive approximation:

$$e^{-\tau(T+V)} \approx e^{-\tau T} e^{-\tau V}. \tag{A.9}$$

The correctness of the approximation in the limit of small τ for the full imaginary time is guaranteed by the Trotter formula [88]:

$$e^{-\beta(T+V)} = \lim_{M\to\infty} \left(e^{-\tau T} e^{-\tau V}\right)^M. \tag{A.10}$$

The primitive approximation in position space reads again as a convolution of two terms:

$$\rho(R_0, R_2, \tau) \approx \int dR_1 \left\langle R_0 \mid e^{-\tau T} \mid R_1 \right\rangle \left\langle R_1 \mid e^{-\tau V} \mid R_2 \right\rangle. \tag{A.11}$$

We can plug this expression for $\rho(R_0, R_2, \tau)$ into Eq. A.8. We are left with finding an expression for the integrand of Eq. A.11. The potential operator is diagonal in position space and it is easily evaluated:

$$\left\langle R_1 \mid e^{-\tau V} \mid R_2 \right\rangle = e^{-\tau V(R_1)} \delta(R_1 - R_2), \tag{A.12}$$

while the kinetic term can be calculated expanding in the eigenfunctions of free particles of mass m in a box of side L in d dimensions in periodic boundary conditions:

$$\begin{aligned} \left\langle R_1 \mid e^{-\tau T} \mid R_2 \right\rangle = \sum_i \left\langle R_1 \mid K_i \right\rangle \left\langle K_i \mid e^{-\tau T} \mid K_i \right\rangle \left\langle K_i \mid R_2 \right\rangle = \\ \sum_i L^{-d} e^{-\tau \frac{\hbar^2}{2m} K_i^2} e^{-iK_i(R_2 - R_1)} \approx (4\pi\lambda\tau)^{-d/2} e^{-\frac{(R_0 - R_1)^2}{4\lambda\tau}} \end{aligned} \tag{A.13}$$

where $\lambda = \hbar^2/2m$. The last term of Eq. A.13 is obtained by substituting the summation by an integral, which is appropriate in the limit of $\lambda\tau \ll L^2$. Combining Eq. A.8, A.11, A.12 and A.13 one obtains the final expression we will be employing for

the unnormalized thermal density matrix:

$$\rho\left(R_0, R_M, \beta\right) = \int dR_1 dR_2 ... dR_{M-1} \left(4\pi\lambda\tau\right)^{-dM/2} exp\left\{-\sum_i^M \left[\frac{\left(R_{i-1} - R_i\right)^2}{4\lambda\tau} + \tau V\left(R_i\right)\right]\right\}. \quad \text{(A.14)}$$

If we substitute Eq. A.14 into A.6 we see that the evaluation of the expectation value of an operator is an highly dimensional integral formally equivalent to the calculation of the ensemble average of a classical system.

$$\langle O \rangle = \frac{\int dR_0 ... dR_M \left(4\pi\lambda\tau\right)^{-dM/2} e^{\left\{-\sum_i^M \left[\frac{\left(R_{i-1} - R_i\right)^2}{4\lambda\tau} + \tau V(R_i)\right]\right\}} \langle R_M \mid O \mid R_0 \rangle}{\int dR_0\, dR_M\, \rho\left(R_0, R_M, \beta\right)}. \quad \text{(A.15)}$$

The convolution representing the density matrix can be viewed as the distribution function, integrating to the partition function. The basis of PIMC is to apply the same Monte Carlo integration techniques used for classical systems to quantum systems, taking advantage of this mapping: namely, to sample $\langle R_M \mid O \mid R_0 \rangle$ according to a probability distribution given by $\rho\left(R', R'', \beta\right)$ expressed in its expanded form. Notice that the latter is a proper unnormalized density distribution for Bosonic systems, being positive at all temperatures. We will not be considering here Fermionic systems, for which this is no longer true and approximations are needed.

If we closer inspect this mapping, we see that the integrand of Eq. A.15 is equivalent to the distribution function of classical particles which are interacting trough $V(R)$ if they pertain to the same time slice, while particles pertaining to neighbouring time slices interact through a spring-like potential $\frac{\left(R_{i-1} - R_i\right)^2}{4\lambda\tau}$. In other words, if we unfold the set $R_i = (r_i^1, r_i^2, ..., r_i^N)$ of particles position we have used in the preceding equations and consider the individual quantum particle position, we see that the position of each quantum particle at different imaginary times are in this mapping classical particles, connected by "springs", and interacting trough the potential term only among the same time slice. In the following, we will call the classical particles "images", to distinguish them from the quantum particle they are mapped from, and the springs connecting the images will be called "links". These links and images form "paths" in configuration space, often called "world lines" in PIMC jargon. Because thermodynamic properties or properties diagonal in space are determined by the trace of the density matrix, in PIMC world lines are closed on themselves. In this way, we are introducing in PIMC a periodicity in imaginary time, imposing the 0^{th} and the M^{th} time slice to be the same. Also, we are making all time slices equivalent as far as

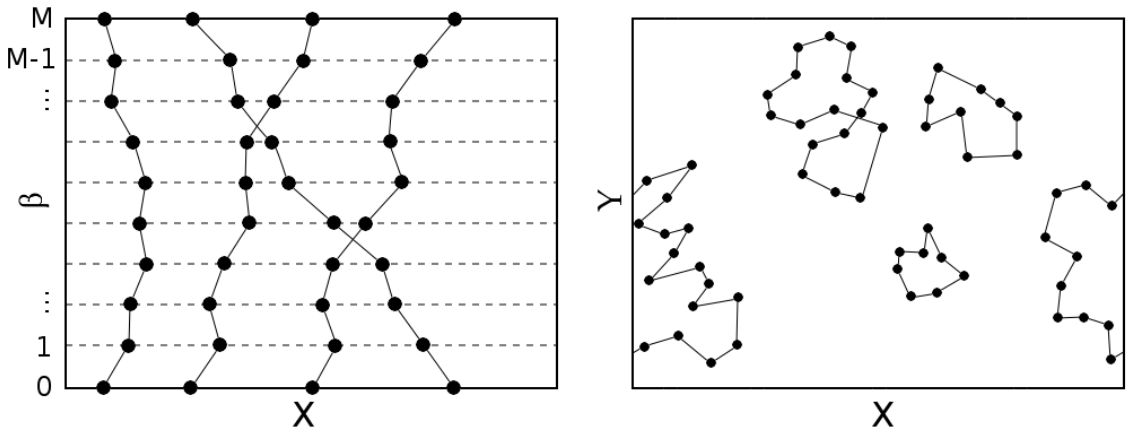

Figure A.1: Time-space representation of a PIMC configuration (left panel), with the diagonality condition $R_0 = R_M$ imposed, and two dimensional spatial representation of permutation cycles, see text. In this figure, full circles represent the position of the images, while the line connecting them is the link, associated with a spring-like term in the distribution function.

their distribution, and therefore the reference time slice index 0 can be set arbitrarily. For distinguishable particles this is equivalent to imposing $R_0 = R_M$, forming closed loops of length M. If however we are dealing with indistinguishable particles, we have to allow all possible permutations of particles to occur, namely $R_0 = P_h\,(R_M)$, where P ia a permutation operator and h a generic permutation index. This means that each world line can either close on itself at time slice M, or close at the beginning of another world line at time 0. Therefore, the loop we are forming can have a length multiple of M, involving permutation of more than one quantum particle. For this reason, world lines are also referred to as "permutation cycles". These permutations need to be sampled, in order to obtain correct estimates for the observables of a Bosonic system. While it is relatively simple to design efficient Metropolis Monte Carlo moves to sample paths configurations at given permutation, it is much harder to sample the permutations themselves. This is because permutations are topologically different, so that there exists no continuous (or "local") transformation in the space of configurations that allows changing them. As a result, designing Metropolis Monte Carlo moves that perform this operation is found too difficult and system size dependent. Eventually, it is found that permutation sampling is exponentially depressed with system size, making permutations "topologically locked" for large sizes. This feature strongly limits the ability of PIMC of extrapolating to the thermodynamic limit, and is the main factor limiting its success. This is exactly the problem

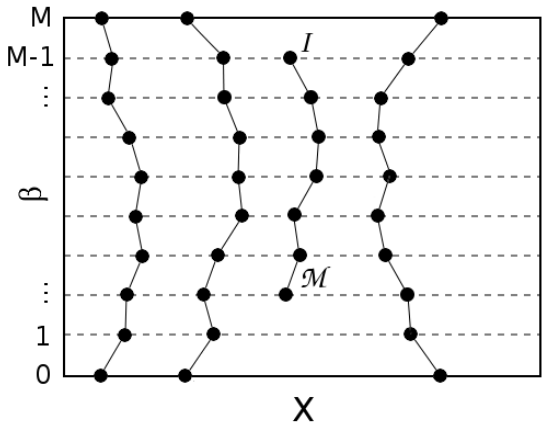

Figure A.2: Time-space representation of a Worm algorithm configuration in which there is an open world line, as well as closed permutation cycles. The head and the tail of the worm are called Ira $\mathcal{I}$ and Masha $\mathcal{M}$.

it is solved by a simple but ingenious modification of the PIMC algorithm called the "worm algorithm", which is described in the following section.

A.2 Worm Algorithm

The extension of PIMC called "worm algorithm" was first derived for lattice models by Prokof'ev [89]. It was subsequently extended to continuous space by Boninsegni [90, 91]. We will be dealing here with the latter version of the algorithm. Worm algorithm not only eliminates the problem of the exponential suppression of permutation sampling with system size, but also makes it easier to calculate off-diagonal correlations, such as single particle particle Matsubara Green function.

The worm algorithm works by expanding the space of allowed configuration of PIMC. In standard path integral, the set of valid configurations is the one in which all world lines form closed permutation cycles[1]. This set of configuration is referred to as the Z sector. In the worm algorithm, configuration in which there is one open world line, called "the worm", are allowed. The set of "open" configurations is called the G sector. An example of such a configuration is given in Figure A.2 This modification is critical for performances, as it eliminates the problem of topological locking of permutations, by allowing cutting and closing of cycles. As a result, in the worm algorithm it is possible to design local Monte Carlo moves able to sample permutations in the

[1]When measuring off-diagonal correlations, open configurations are admitted also in PIMC, but are not used for changing permutations.

Z sector by moving through the G sector. Most of the estimators are calculated in the Z sector, although the G sector is key to the calculation of off-diagonal imaginary time correlations such as the Matsubara Green function. In order to enter into the details of the algorithm, it is useful to introduce some standard nomenclature. We will denote with M the number of time slices, N_{im} the number of images in a certain configuration, while with N we denote the number of quantum particles present in a configuration pertaining to the Z sector. Note that in the this sector the number of links is equal to N_{im}, while in the G sector is $N_{im} - 1$, because we have to remove at least a link in order to have an open configuration. Following reference [90], the head and the tail of the worm are called Ira $\mathcal{I}$ and Masha $\mathcal{M}$. In the following, we will sometime refer to specific images with Greek letters, such as α. The position of an image at time slice i and image index j will be denoted by r_i^j, or $r(\alpha)$ if we are referring to an already named image. The vector containing all images coordinates at time slice i is R_i. We will denote as $next(\alpha)$ the image connected to alpha by a link and forward in time, while with $prev(\alpha)$ the preceding image. The $next(next(\alpha))$ will be written as $next^2(\alpha)$, and similarly for the $prev$ operator. Sometime we will sum over time slice indexes; as these are periodic, the result is to be taken modulus M. The same applies also for difference in time slices.

As there is the need of removing and adding part of world lines, the worm algorithm is most easily defined in the grand canonical ensemble, although it is not difficult do design a canonical version. A valid grand canonical expression of the sampling distribution, Eq. A.14, is:

$$\pi(R_0, R_M, \beta) = \int dR_1 ... dR_{M-1} (4\pi\lambda\tau)^{-dM/2} e^{-\sum_i^M \left[\frac{(R_{i-1}-R_i)^2}{4\lambda\tau} + \tau(V(R_i)-\mu N_i)\right]}, \quad \text{(A.16)}$$

where N_i is the number of images at time slice index i. This expression reduces to the grand canonical expression for the unnormalized density matrix in the Z sector. This distribution is sampled with a generalized Metropolis-Hastings algorithm [92, 93]. We recall that in this algorithm the probability of acceptance of a new world line configuration s' starting from configuration s is:

$$A(s, s') = min\left\{1, \frac{T(s', s)\,\Pi(s')}{T(s, s')\,\Pi(s)}\right\}, \quad \text{(A.17)}$$

where $T(s, s')$ is the probability of proposing a transition from s to s', and $\Pi(s)$ is the distribution function A.16. In the following, we briefly describe the Monte Carlo Metropolis moves constructed according to Eq. A.17 implemented in the code

employed in this thesis. A graphical illustration of the effect of these MC moves on the configuration of world lines is provided in Figure A.3.

Replace The replace MC moves applies both to the G and to the Z sector, and is the typical MC move one would find in standard PIMC. An image α is chosen at random. It is traced forward in imaginary time for N_{rp} time steps, individuating a new particle we will denote by $\upsilon = next^{N_{rp}}(\alpha)$. If in this process $\mathcal{I}$ is encountered, or if $\upsilon = \mathcal{I}$, the move is rejected. N_{rp} is a uniformly distributed integer number $\in [1, N_{max}]$, where N_{max} is an algorithm parameter. A new piece of world line connecting α to υ is sampled from the product of N_{rp} free particle propagators $\prod_{i=1}^{N_{rp}} \rho_0(r_{i-1}, r_i, \tau)$.

$$\rho_0(r_\alpha, r_\omega, \tau) = (4\pi\lambda\tau)^{-dM/2} exp\left\{\frac{(r_{i-1} - r_i)^2}{4\lambda\tau}\right\}. \tag{A.18}$$

The new piece of the world line is substituted to the original one with acceptance probability:

$$A_{rp} = min\left\{1, e^{-\Delta U}\right\}, \tag{A.19}$$

where $\Delta U = \sum_i V(R_i(s')) - V(R_i(s))$ is the difference in total potential energy between initial and final configurations.

Open The open MC move brings the system from the Z sector to the G sector by opening a permutation cycle. It starts by checking that the system is in the Z sector, otherwise the move is rejected. A random integer N_{op} uniformly distributed in the interval $[1, N_{max}]$ is selected. An image α is selected at random, and $N_{op} - 1$ images and N_{op} links following the selected image are removed, so that α becomes the head $\mathcal{I}$ of the worm and the N_{op}^{th} image after α becomes the worm tail $\mathcal{M}$. The probability of acceptance of a proposed move generated in this way is:

$$A_{op} = min\left\{1, \frac{CN_{max}N_{im}e^{-\Delta U - \mu N_{op}\tau}}{\rho_0(r_\mathcal{I}, r_\mathcal{M}, N_{op}\tau)}\right\}, \tag{A.20}$$

where N_{im} is the number of images in the starting configuration, in this case equal to MN. C is an algorithm constant determining the relative statistical weight of Z and G sectors, $\Delta U = \sum_i V(R_i(s')) - V(R_i(s))$ is the difference in total potential energy between configurations.

Close The close MC move is the reverse of the open move. This move is applicable in the G sector only, and if accepted brings the system into the Z sector. It starts by checking the difference N_{cl} in time slices between $\mathcal{M}$ and $\mathcal{I}$. If $N_{cl} \notin [1, N_{max}]$ the move is rejected, otherwise a portion of path connecting $\mathcal{I}$ and $\mathcal{M}$ is proposed. This path $\{r_{\mathcal{I}}, r_1, ..., r_{N_{cl}-1}, r_{\mathcal{M}}\}$, with $r_0 = r_{\mathcal{I}}$ and $r_{N_{cl}} = r_{\mathcal{M}}$, is sampled from the product of N_{cl} free particle propagators $\prod_{i=1}^{N_{cl}} \rho_0 (r_{i-1}, r_i, \tau)$. The probability of acceptance of the proposed move is then:

$$A_{cl} = min \left\{ 1, \frac{e^{-\Delta U + \mu N_{cl} \tau} \rho_0 (r_{\mathcal{I}}, r_{\mathcal{M}}, N_{cl}\tau)}{C N_{max} N_{im}} \right\}, \tag{A.21}$$

where N_{im} is the number of images in the final configuration, and ΔU is again the difference in potential energy between final and initial configurations. Open and close moves are conjugated, in the sense that the probability of acceptance of each depends mutually on the way the proposal of the other is designed. In other words, in Eq. A.17, if $T(s, s')$ is the probability of proposing a certain open move, $T(s', s)$ is the probability of proposing the opposite close move.

Insert The insert MC move brings the system from the Z sector to the G sector by inserting a new worm in the system. If the system is in the Z sector, we select a position in space and a time slice index at random. Then we generate new $N_{in} \in [1, N_{max}]$ images distributed according to $N_{in} - 1$ free particles propagators. The move is then accepted with probability

$$A_{in} = min \left\{ 1, C N_{max} V M e^{-\Delta U + \mu N_{in} \tau} \right\}, \tag{A.22}$$

where V is the volume of the system.

Remove The remove MC move is the conjugate of the insert move. It brings the system from the G sector to the Z sector by attempting to remove the worm. If the system is in the G sector, it is checked the length of the worm N_{re}, defined as the number of images it is composed of. If this length $N_{re} \in [1, N_{max}]$, then the worm is removed with probability

$$A_{re} = min \left\{ 1, \frac{e^{-\Delta U - \mu N_{re} \tau}}{C N_{max} V M} \right\}. \tag{A.23}$$

Advance The advance MC move applies to the G sector and keeps the system in it. A new piece of word line of length of $N_{in} \in [1, N_{max}]$ images is generated, distributed

according to N_{ad} free particles propagators, starting from $\mathcal{I}$. The new piece of world line is inserted in the system with probability

$$A_{ad} = min\left\{1, e^{-\Delta U + \mu N_{ad}\tau}\right\}. \tag{A.24}$$

The new $\mathcal{I}$ is now the end of the added part of world line.

Recede The recede MC move is the conjugate of the advance move, and applies to the G sector. If the system is in the G sector, a number $N_{rd} \in [1, N_{max}]$ is selected. The length of the worm by number of images is checked, if this is smaller or equal to N_{re} the move is rejected. Otherwise N_{rd} images are removed from the worm, starting from $\mathcal{I}$ and going backwards in imaginary time, with probability

$$A_{rd} = min\left\{1, e^{-\Delta U - \mu N_{rd}\tau}\right\}. \tag{A.25}$$

Swap The swap move is defined in the G sector, and is the critical move for efficient permutation sampling (although also the previously described moves have the ability to change the permutations). After checking whether the system is in the G sector, than the time slice index of $\mathcal{I}$, denoted by j, is identified. The $j + N_{sw}$ time slice is considered, $N_{sw} \in [1, M-1]$ being an algorithm parameter and an image α pertaining to this time slice is chosen according to the probability distribution

$$\begin{aligned} \pi(\alpha) &= \frac{\rho_0\left(r_{\mathcal{I}}, r(\alpha), N_{sw}\tau\right)}{\varsigma\left(\mathcal{I}, N_{sw}\right)} \\ \varsigma\left(\mathcal{I}, N_{sw}\right) &= \sum_i \rho_0\left(r_{\mathcal{I}}, r^i_{j+N_{sw}}, N_{sw}\tau\right), \end{aligned} \tag{A.26}$$

where the index i in the summation spans over all the images at time slice $j + N_{sw}$. Now we trace back starting from α for N_{sw} time slices, individuating an image we will denote with $\iota = prev^{N_{sw}}(\alpha)$. If when tracing back from α, $\mathcal{M}$ is encountered, or if $\iota = \mathcal{M}$, the move is rejected. The quantity $\varsigma(\iota, N_{sw})$ is calculated, in the same way as $\varsigma(\mathcal{I}, N_{sw})$. Now, a new piece of word line constituted by $N_{sw} - 1$ images is constructed, sampled as the product of N_{sw} free particles propagators connecting $\mathcal{I}$ to α. Then, with probability

$$A_{sw} = min\left\{1, \frac{e^{-\Delta U}\varsigma\left(\mathcal{I}, N_{sw}\right)}{\varsigma\left(\iota, N_{sw}\right)}\right\}, \tag{A.27}$$

the world line connecting ι to α is removed and substituted with the proposed one connecting $\mathcal{I}$ to α. In this case ι becomes the new $\mathcal{I}$.

The code is written in Fortran90, allowing for both canonical and grand-canonical ensemble. A number of estimators such as potential and kinetic energy, intermediate scattering function [86], superfluid fraction [94], pair distribution function, one body density matrix, Green function [90], are included in the code. We explicitly report here only the expression for the estimator of the superfuid density in periodic boundary conditions, as this quantity is crucial in our analysis.

$$\rho_s = \frac{m \langle \mathbf{W} \rangle^2}{\hbar^2 \beta N}, \tag{A.28}$$

where $\mathbf{W}$ is the winding number estimator, defined as

$$\mathbf{W} = \sum_{i,t} \mathbf{r}_i(t+1) - \mathbf{r}_i(t), \tag{A.29}$$

where i is the particle index and t the time slice index. The distance $\mathbf{r}_i(t+1) - \mathbf{r}_i(t)$ has to be calculated using the periodic boundary replica of image $i+1$ nearest to image i. The sum in Eq. A.29 can be different from zero if some permutation cycle winds around the periodic boundary condition, as paths that never cross the boundary do not contribute into the summation.

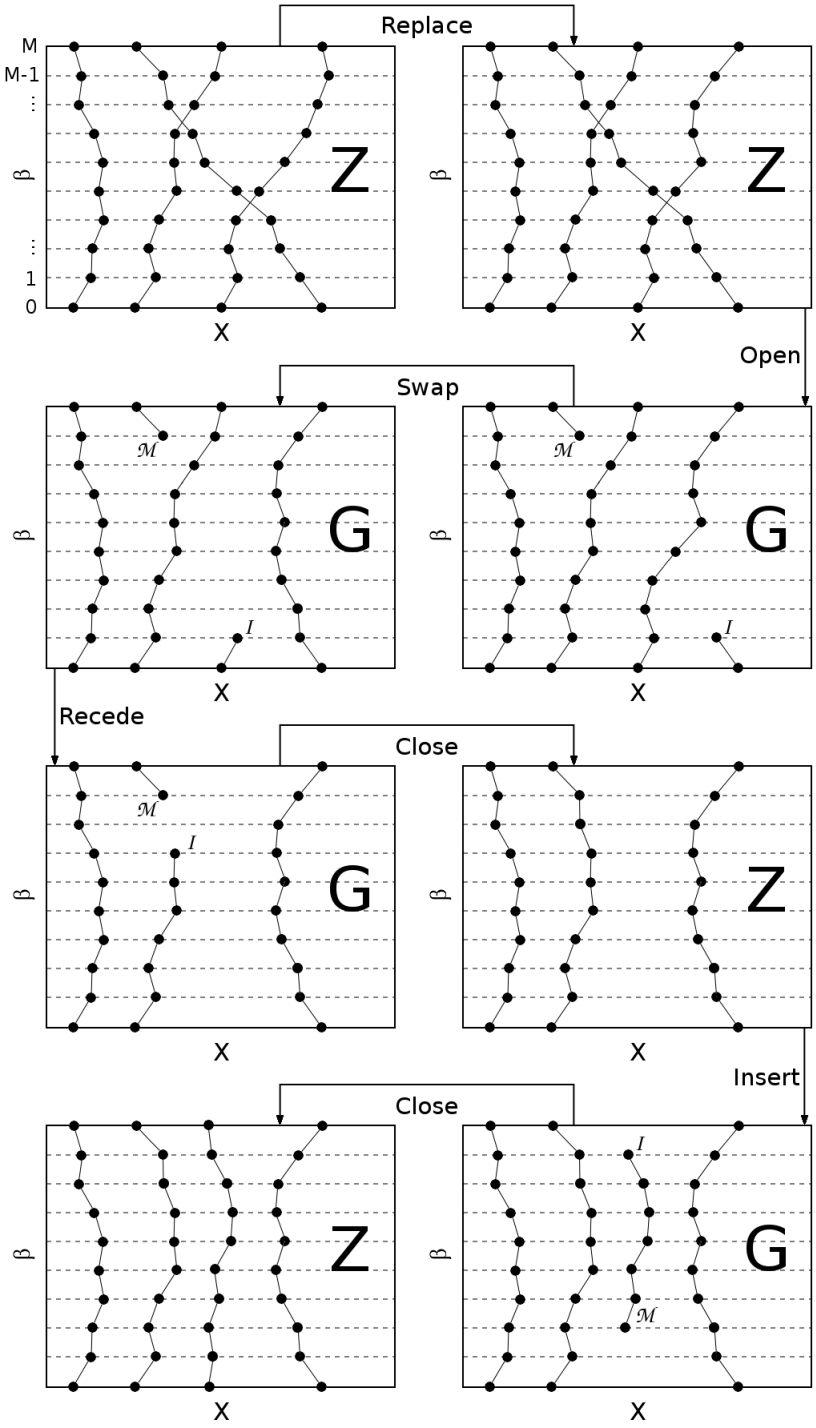

Figure A.3: Illustration of the effect of the various Metropolis Monte Carlo moves on the configuration of the world lines in the worm algorithm.

Appendix B

B.1 Variational Monte Carlo

Variational Monte Carlo (VMC) is a way of calculating stochastically expectation values of a trial wavefunction ψ_T. It is based on the observation that any expectation value can be expressed as

$$\begin{gathered}\langle O\rangle = \frac{\langle \psi_T \mid O \mid \psi_T\rangle}{\langle \psi_T \mid \psi_T\rangle} = \frac{\int \mathrm{d}\mathbf{r}\, \langle \psi_T \mid \mathbf{r}\rangle \langle \mathbf{r} \mid \psi_T\rangle \frac{\langle \mathbf{r}|O|\psi_T\rangle}{\langle \mathbf{r}|\psi_T\rangle}}{\int \mathrm{d}\mathbf{r}\, \langle \psi_T \mid \mathbf{r}\rangle \langle \mathbf{r} \mid \psi_T\rangle} = \frac{\int \mathrm{d}\mathbf{r}\, \Pi(\mathbf{r},\psi_T) E_O(\mathbf{r})}{\int \mathrm{d}\mathbf{r}\, \Pi(\mathbf{r},\psi_T)} \\ \Pi(\mathbf{r},\psi_T) = |\psi_T(\mathbf{r})|^2; \quad E_O(\mathbf{r},\psi_T) = \frac{[O\psi_T](\mathbf{r})}{\psi_T(\mathbf{r})},\end{gathered} \tag{B.1}$$

where $\mathbf{r}$ denotes particle coordinates. In case the operator O of our interest is the Hamiltonian of the system $\mathcal{H}$, the expectation value is the total energy of the system E and the term E_O takes the name "local energy"

$$E_l(\mathbf{r},\psi_T) = \frac{[\mathcal{H}\psi_T](\mathbf{r})}{\psi_T(\mathbf{r})}. \tag{B.2}$$

Eq. B.1 is in general a highly multidimensional integral. Moreover, the term $\Pi(\mathbf{r},\psi_T)$ as a function of $\mathbf{r}$ can be interpreted as an unnormalized probability, as it is positive definite. These two observation point to Monte Carlo integration as a way to solve Eq. B.1. By sampling $\mathbf{r}$ distributed according to $\Pi(\mathbf{r},\psi_T)$, the estimate for the total energy E is evaluated as the average $E_l(\mathbf{r})$ over this sampling. The technique of choice to perform the sampling is most times the Metropolis-Hastings algorithm [92,93], Eq. A.17 (where s and s' have to be interpreted here as $\mathbf{r}$ and $\mathbf{r}'$). A possible proposal probability for $T(\mathbf{r}',\mathbf{r})$ is a multidimensional gaussian distribution centered on $\mathbf{r}$, of suitable amplitude in order to obtain a reasonable acceptance ratio, but also more sophisticated strategies can be applied if one aims at high efficiency [95]. Eq. B.1 for the total energy retains its variational property in the stochastic integration:

$\langle E \rangle \geq E_g$, where E_g is the true ground state energy of the system, and $\langle E \rangle = E_g$ only if $\psi_T = \psi_g$. VMC also exploits the important "zero variance" property of the local energy, which states that if our trial wave function is equal to the true ground state wavefunction $\psi_T = \psi_g$, the local energy will be identically equal to the ground state energy $E_l(\mathbf{r}, \psi_T) = E_g$, as it is easily verified from Eq. B.2. This properties guarantees that by improving the trial wave function ψ_T the variance of $E_l(\mathbf{r}, \psi_T)$ sampled over $\Pi(\mathbf{r}, \psi_T)$ will decrease, making the stochastic integration cheaper. The advantage of VMC with respect to other wavefunction based methods is in that we need not to be able to perform explicit integrations over $\psi_T(\mathbf{r})$, therefore we need not to use separable wavefunctions, say as a series of Slater determinants. In fact, the only requisite we have is to be able to evaluate $\psi_T(\mathbf{r})$ and its derivatives, possibly in a fast way in order to make the computation efficient. Therefore, we have the possibility of including explicit correlations into ψ_T, e.g. by the use of a Jastrow term, depending on all the electron and nuclear coordinates $\mathbf{r}$ and $\mathbf{R}$ simultaneously $J(\mathbf{r}, \mathbf{R})$, which is typically a positive definite function expressed as an exponential of some correlation functions. In general there is great freedom in the choice of the form of these functions, which may even include high order correlation terms. As the Jastrow term has the ability to recover much of the correlation energy, for many applications there may be no need to include correlation through the use of multi-determinantal wavefunctions, therefore a common VMC wavefunction is the so called Slater-Jastrow, constituted by single Slater determinant multiplied by a Jastrow factor.

B.2 Diffusion Monte Carlo

Diffusion Monte Carlo is a method aimed at stochastically projecting a trial wave functions toward the true ground state wavefunction. We start from the Schrödinger equation

$$i\hbar \frac{\partial \psi(\mathbf{r}, t)}{\partial t} = (\mathcal{H} - E_t)\psi(\mathbf{r}, t) = -\frac{\hbar^2}{2m} \bigtriangledown^2 \psi(\mathbf{r}, t) + (V - E_t)\psi(\mathbf{r}, t), \tag{B.3}$$

where E_t is he so called "trial energy", providing a rigid shift of the eigenvalues of the equation. Now we consider the imaginary-time evolution of Eq. B.3, starting from an initial trial wavefunction $\psi_T(\mathbf{r})$ at time zero

$$\psi(\mathbf{r}, t) = exp\left[-t\left(\mathcal{H} - E_t\right)\right] \psi_T(\mathbf{r}). \tag{B.4}$$

In the limit $t \to \infty$ the procedure will exponentially project out the eigenfunction not perpendicular to $\psi_g(\mathbf{r})$ corresponding to the lowest eigenvalue, as can be seen by decomposing $\psi_T(\mathbf{r})$ in eigenfunctions of $\mathcal{H}$. Eq. B.4 clarifies the role of E_t as a constant needed to keep the normalization of the lowest eigenvalue. Eq. B.4 can also be rewritten in a convolution form:

$$\begin{aligned} \psi(\mathbf{r}', t+\tau) &= \int \mathrm{d}\mathbf{r}\, G(\mathbf{r}, \mathbf{r}', \tau)\psi(\mathbf{r}, t) \\ G(\mathbf{r}, \mathbf{r}', \tau) &= \langle \mathbf{r}' \mid exp\left[-\tau\left(\mathcal{H} - E_t\right)\right] \mid \mathbf{r}\rangle . \end{aligned} \tag{B.5}$$

A short time approximation for the Green function $G(\mathbf{r}, \mathbf{r}', \tau)$ can be devised, in a way completely analogous to the procedure described in section A.1 for the short time approximation of $\rho(R, R', \beta)$. By using the Trotter decomposition one obtains:

$$G(\mathbf{r}, \mathbf{r}', \tau) \approx (4\pi\lambda\tau)^{-3N/2} exp\left[-\frac{(\mathbf{r}-\mathbf{r}')^2}{2\tau}\right] exp\left[-\frac{\tau}{2}\left(V(\mathbf{r}) + V(\mathbf{r}') - 2E_t\right)\right], \tag{B.6}$$

where $V(\mathbf{r})$ is the potential energy operator of the Hamiltonian, $\lambda = \frac{\hbar^2}{2m}$ and N the number of particles. In principle, with repeated application of $G(\mathbf{r}, \mathbf{r}', \tau)$, one can sample the true ground state wavefunction of the system: it is enough to interpret Eq. B.5 as a master equation of a stochastic process. We can select a position $\mathbf{r}$ distributed according to the initial $\psi_T(\mathbf{r})$, and let its position evolve stochastically by repeatedly applying $G(\mathbf{r}, \mathbf{r}', \tau)$ until at imaginary time t all the components other from the ground state are projected out. This procedure can be repeated at will, obtaining sampling points at time t distributed according ψ_g. Minor complications arise from the renormalizing factor $exp\left[-\frac{\tau}{2}\left(V(\mathbf{r}) + V(\mathbf{r}') - 2E_t\right)\right]$ which can be dealt with in a stochastic approach either by assigning a weight to the sampling points, also referred to as "walkers", by the branching approach (see below). Therefore in principle, by this procedure, one can sample the true ground state wavefunction of the system. This can then be used to estimate expectation values, such as the energy, in the same spirit of VMC, by a weighted average of the local energy over this sampling.

However, here the well known sign problem comes into play: if we are dealing with a Fermionic system, the initial $\psi_T(\mathbf{r})$ will in general not be positive definite, therefore is not a proper probability distribution. If we tackle this problem by distributing the initial walkers as $|\psi_T(\mathbf{r})|$, and keeping track of the sing in the walker's weight, cancellation of weights will exponentially increase the statistical uncertainty of any expectation value we may wish to calculate on the sampled distribution. This is easily

understood if we imagine to separate the positive and negative parts of ψ^+ and ψ^-, and project them separately, taking advantage of the linearity of Eq.B.4. Both the positive and negative population will distribute toward the Bosonic ground state at $t \to \infty$, canceling each other out. In order to solve this notorious problem, one of the most used approaches is the so-called fixed-node approximation. As the name suggests, it is a modification of the imaginary time propagation devised so to prevent walkers from crossing the nodes, i.e. hypersurfaces in which $\psi_T(\mathbf{r})$=0. By doing so one introduces an approximation as it is fixing the nodal surface, therefore the projection operation cannot reach the true ground state wavefunction anymore. This constraint can be introduced into the algorithm by means of an importance sampling procedure. We start by modifying the Schrödinger equation B.3 by multiplying both sides by $\psi_T(\mathbf{r})$, obtaining an equation for $f(\mathbf{r},t) = \psi(\mathbf{r},t)\,\psi_T(\mathbf{r})$

$$\begin{aligned} i\hbar\frac{\partial f(\mathbf{r},t)}{\partial t} &= -\frac{\hbar^2}{2m}\bigtriangledown^2 f(\mathbf{r},t) + \frac{\hbar^2}{2m}\bigtriangledown\cdot(\mathbf{v}_d(\mathbf{r})f(\mathbf{r},t)) + (E_l(\mathbf{r}) - E_t)\,f(\mathbf{r},t) \\ E_l(\mathbf{r},\psi_T) &= \frac{[\mathcal{H}\psi_T](\mathbf{r})}{\psi_T(\mathbf{r})} \\ \mathbf{v}_d(\mathbf{r}) &= \bigtriangledown ln\,|\psi_T(\mathbf{r})|\,, \end{aligned} \tag{B.7}$$

where E_l is the local energy and $\mathbf{v}_d(\mathbf{r})$ is the "drift velocity" term. This equation is not equivalent to the original Schrödinger equation, as the function $\psi_T(\mathbf{r})$ is zero at the nodes, and we are effectively decoupling the evolution of ψ in different nodal pockets, i.e. regions enclosed into a nodal surface. In analogy to what done for the original evolution in imaginary time, also for the evolution driven by Eq. B.7 can be written as a master equation of a stochastic process:

$$f(\mathbf{r}',t+\tau) = \int \mathrm{d}\mathbf{r}\; G(\mathbf{r},\mathbf{r}',\tau) f(\mathbf{r},t)\,. \tag{B.8}$$

We can write by the Trotter decomposition a short time approximation (exact in the limit of $\tau \to 0$) of the Green function as:

$$\begin{aligned} G(\mathbf{r},\mathbf{r}',\tau) &= \int d\mathbf{r}''\,d\mathbf{r}'''G_{g/d}(\mathbf{r}''',\mathbf{r}',\tau)\,G_{diff}(\mathbf{r}'',\mathbf{r}''',\tau)\,G_{drift}(\mathbf{r},\mathbf{r}'',\tau) \\ G_{drift}(\mathbf{r},\mathbf{r}',\tau) &= (1-\tau\bigtriangledown\cdot\mathbf{v}_d(\mathbf{r}))\,\delta(\mathbf{r}'-\mathbf{r}-\tau\mathbf{v}_d(\mathbf{r})) \\ G_{diff}(\mathbf{r},\mathbf{r}',\tau) &= (4\pi\lambda\tau)^{-3N/2}\,exp\left[-\frac{(\mathbf{r}-\mathbf{r}')^2}{2\tau}\right] \\ G_{g/d}(\mathbf{r},\mathbf{r}',\tau) &= exp\left[-\tau\left(E_l(\mathbf{r},\psi_T)-E_t\right)\right]\delta(\mathbf{r}-\mathbf{r}')\,. \end{aligned} \tag{B.9}$$

The three terms of the Green function have the following interpretation in the stochastic implementation:

G_{drift} Drift term. Has the effect of an importance sampling: it drives walkers toward the region in which $|\psi_T(\mathbf{r})|$ is larger. It also ensures, in the limit of $\tau \to 0$, that walkers do not cross the nodal surface, where the term $\mathbf{v}_d(\mathbf{r})$ diverges.

G_{diff} Diffusion term. It diffuses walkers by randomly displacing them of a normally distributed quantity.

$G_{g/d}$ Growth/decay term. It is a statistical weight acquired by the walker. This term can be dealt with in a branching approach (see below).

We will also sometime refer to the convolution of G_{drift} and G_{diff} as the drift-diffusion term:

$$G_{dd}(\mathbf{r}, \mathbf{r}', \tau) = \int d\mathbf{r}'' G_{diff}\left(\mathbf{r}'', \mathbf{r}', \tau\right) G_{drift}\left(\mathbf{r}, \mathbf{r}'', \tau\right). \tag{B.10}$$

The procedure then follows as in the original dynamic. Starting from a random walker distributed initially as $|\psi_T(\mathbf{r})|^2$, it is projected in time by successive small time steps τ until a stationary distribution $f_0(\mathbf{r}) = \psi_0(\mathbf{r})\psi_T(\mathbf{r})$ is reached. Naturally results are to be extrapolated in the $\tau \to 0$ limit. This distribution can be used for computing expectation values such as the energy, trough the mixed estimator

$$\begin{gathered} \langle E \rangle = \frac{\langle \psi_0 \mid \mathcal{H} \mid \psi_T \rangle}{\langle \psi_0 \mid \psi_T \rangle} = \frac{\int \mathrm{d}\mathbf{r}\, \Pi\left(\mathbf{r}, \psi_T\right) E_l\left(\mathbf{r}, \psi_T\right)}{\int \mathrm{d}\mathbf{r}\, \Pi\left(\mathbf{r}, \psi_T\right)} \\ \Pi\left(\mathbf{r}, \psi_T\right) = f_0\left(\mathbf{r}\right). \end{gathered} \tag{B.11}$$

The algorithm retains the variational property in that $\langle E \rangle \geq E_g$. Differently with VMC there is a less strict condition for the two quantities to be equal, $\langle E \rangle = E_g$ only if the nodal surface of ψ_T is equal to the one of ψ_g. It can be in fact demonstrated that the expectation value of the energy is a function on the nodal surface of ψ_T only. As a consequence, a Jastrow factor we might be adding into ψ_T does not have any effect on $\langle E \rangle$. However, the inclusion of a Jastrow factor is critical for an efficient implementation of the algorithm, as the growth/decay Green function fluctuations term depends on the details of ψ_T. This can be seen easily, for example if $\psi_T = \psi_g$ then $E_l = E_g$, i.e. as in VMC the zero variance property holds, with the $G_{g/d}$ going to a constant.

The algorithm as it is presented up to now lacks some key ingredients for an efficient implementation, see e.g. [96]. We will not enter the details of these technicalities,

being a description of these algorithm beyond the scope of this thesis, but we will mention the most relevant. First, it is fundamental to devise a strategy to control the weight fluctuations of walker, due to the $G_{g/d}$ term. The most common solution is represented by the many walker formulation with branching. Instead of using a single walker, we employ a population of them propagating simultaneously. After each time step, i.e. after each application of $G(\mathbf{r}, \mathbf{r}', \tau)$, each walker will have a weight multiplied by the $G_{g/d}(\mathbf{r}, \mathbf{r}', \tau)$. Instead of simply keeping track of the weight for each single walker, walkers with high weight are replicated in copies, while walkers with low weight are removed. This proves to be a much more efficient methodology for large many-body systems, as we are avoiding to propagate walkers that have small statistical weight. Once this is done, one has to guarantee that the population of these walkers does not diverge or vanish exponentially, by acting on the constant E_t, entering in the $G_{g/d}$. However, by dynamically acting on E_t one introduces a bias, as it is suppressing weight fluctuations in the population, which would be present even if $E_t =< E >$. Weight fluctuations are then re-introduced back into averages by properly reweighting the walkers. Other problems are arising at the nodes of ψ_T, where E_l and $\mathbf{v}_d$ diverges. For this reason near the nodes G_{drift} of Eq. B.9 is a rather poor approximation of the true Green function, so that better approximations have been devised. At finite τ the density of walkers near a node is not zero, due to the approximation in the Green function. This problem can be tackled by adding an accept/reject step that corrects for the imperfect sampling of the G_{drift} and G_{diff} terms. In practice, after the drift-diffusion term $G_{dd}(\mathbf{r}, \mathbf{r}', \tau)$ is applied, the new position $\mathbf{r}'$ is accepted with probability:

$$P = min\left\{1, \frac{|\psi_T(\mathbf{r}')|^2 G_{dd}(\mathbf{r}, \mathbf{r}', \tau)}{|\psi_T(\mathbf{r})|^2 G_{dd}(\mathbf{r}', r, \tau)}\right\}. \tag{B.12}$$

This Metropolis-Hastings term ensures that the sampling of $G_{dd}(\mathbf{r}, \mathbf{r}', \tau)$ in the limit of $\psi_T = \psi_g$ is performed with no time step error.
Finally, at electron-electron overlap and at electron-nucleus overlap E_l diverges unless proper cusp conditions are satisfied by ψ_T.

Bibliography

[1] E. P. Gross. . *Physical Review*, 106:161, 1957.

[2] A. F. Andreev and I. M. Lifshitz. No Title. *Sov. Phys. JETP*, 29:1107, 1969.

[3] D. J. Thouless. The flow of a dense superfluid. *Annals of Physics*, 52(3):403, 1969.

[4] A. J. Leggett. Can a Solid Be" Superfluid"? *Physical Review Letters*, 25(22):1543, 1970.

[5] G. V. Chester. Speculations on Bose-Einstein condensation and quantum crystals. *Physical Review A*, 2(1):256, 1970.

[6] E. Kim and M. H. W. Chan. Probable observation of a supersolid helium phase. *Nature*, 427(6971):225–7, January 2004.

[7] N. V. Prokofev. What makes a crystal supersolid? *Advances in Physics*, 56(2):381, 2007.

[8] L. Pollet, M. Boninsegni, a. Kuklov, N. Prokofev, B. Svistunov, and M. Troyer. Local Stress and Superfluid Properties of Solid He4. *Physical Review Letters*, 101(9):097202, August 2008.

[9] M. Boninsegni, A. Kuklov, L. Pollet, N. Prokofev, B. Svistunov, and M. Troyer. Fate of Vacancy-Induced Supersolidity in He4. *Physical Review Letters*, 97(8):1–4, August 2006.

[10] D. Kim and M. H. W. Chan. Absence of Supersolidity in Solid Helium in Porous Vycor Glass. *Physical Review Letters*, 109(15):1–5, October 2012.

[11] N. Henkel, R. Nath, and T. Pohl. Three-Dimensional Roton Excitations and Supersolid Formation in Rydberg-Excited Bose-Einstein Condensates. *Physical Review Letters*, 104(19):1–4, May 2010.

[12] F. Cinti, P. Jain, M. Boninsegni, A. Micheli, P. Zoller, and G. Pupillo. Supersolid droplet crystal in a dipole-blockaded gas. *Physical Review Letters*, 105(1):1–4, 2010.

[13] S. Saccani, S. Moroni, and M. Boninsegni. Phase diagram of soft-core bosons in two dimensions. *Physical Review B*, 83(9):092506, March 2011.

[14] S. Saccani, S. Moroni, E. Vitali, and M. Boninsegni. Bose soft discs: a minimal model for supersolidity. *Molecular Physics*, (January 2012):37–41, 2011.

[15] S. Saccani, S. Moroni, and M. Boninsegni. Excitation spectrum of a supersolid. *Physical review letters*, 108(17):175301, April 2012.

[16] B. M. Mladek, D. Gottwald, G. Kahl, M. Neumann, and C. Likos. Formation of Polymorphic Cluster Phases for a Class of Models of Purely Repulsive Soft Spheres. *Physical Review Letters*, 96(4):3–6, January 2006.

[17] C. N. Likos, A. Lang, M. Watzlawek, and H. Löwen. Criterion for determining clustering versus reentrant melting behavior for bounded interaction potentials. *Physical Review E*, 63(031206):031206, 2001.

[18] T. Neuhaus and C. N. Likos. Phonon dispersions of cluster crystals. *Journal of physics. Condensed Matter*, 23(23):234112, July 2011.

[19] J. M. Kosterlitz and D. J. Thouless. *Progress in Low Temperature Physics*, volume 7. 1978.

[20] D. Galli, L. Reatto, and W. Saslow. Bounds for the superfluid fraction from exact quantum Monte Carlo local densities. *Physical Review B*, 76(5):052503, August 2007.

[21] S. Trotzky, L. Pollet, F. Gerbier, U. Schnorrberger, I. Bloch, N. V. Prokofev, B. Svistunov, and M. Troyer. Suppression of the critical temperature for superfluidity near the Mott transition. *Nature Physics*, 6(12):998–1004, October 2010.

[22] D. Stamper-Kurn, A. Chikkatur, A. Görlitz, S. Inouye, S. Gupta, D. Pritchard, and W. Ketterle. Excitation of Phonons in a Bose-Einstein Condensate by Light Scattering. *Physical Review Letters*, 83(15):2876–2879, October 1999.

[23] D. Clément, N. Fabbri, L. Fallani, C. Fort, and M. Inguscio. Bragg Spectroscopy of Strongly Correlated Bosons inOptical Lattices. *Journal of Low Temperature Physics*, 158(1-2):5–15, October 2009.

[24] E. Vitali, M. Rossi, L. Reatto, and D. Galli. Ab initio low-energy dynamics of superfluid and solid He. *Physical Review B*, 82(17):174510, November 2010.

[25] M. Jarrell and J. E. Gubernatis. Bayesian inference and the analytic continuation of imaginary-time quantum Monte Carlo data. *Physics Reports*, 269(3):133–195, May 1996.

[26] M. Boninsegni. Density Fluctuations in Liquid 4He. Path Integrals and Maximum Entropy. *Journal of Low Temperature Physics*, 104:339, 1996.

[27] S. Baroni and S. Moroni. Reptation Quantum Monte Carlo: A Method for Unbiased Ground-State Averages and Imaginary-Time Correlations. *Physical Review Letters*, 82(24):4745–4748, June 1999.

[28] C. Menotti and N. Trivedi. Spectral weight redistribution in strongly correlated bosons in optical lattices. *Physical Review B*, 77(23):1–13, June 2008.

[29] B. Capogrosso-Sansone, . Söyler, N. V. Prokofev, and B. V. Svistunov. Monte Carlo study of the two-dimensional Bose-Hubbard model. *Physical Review A*, 77(1):1–4, January 2008.

[30] C. D. Yoo and A. T. Dorsey. Hydrodynamic theory of supersolids: Variational principle, effective Lagrangian, and density-density correlation function. *Physical Review B*, 81(13):134518, April 2010.

[31] J. Ye. Elementary excitation in a supersolid. *EPL (Europhysics Letters)*, 82(1):16001, April 2008.

[32] C. Josserand, Y. Pomeau, and S. Rica. Coexistence of Ordinary Elasticity and Superfluidity in a Model of a Defect-Free Supersolid. *Physical Review Letters*, 98(19):3–6, May 2007.

[33] J. A. Tarvin, F. Vidal, and T. J. Greytak. Measurements of the dynamic structure factor near the lambda temperature in liquid helium. *Physical Review B*, 15(9):4193, 1977.

[34] H. Eyring. The Activated Complex in Chemical Reactions. *The Journal of Chemical Physics*, 3(2):107, 1935.

[35] E. Wigner. The transition state method. *Transactions of the Faraday Society*, 34:29, 1938.

[36] J. C. Keck. Variational theory of reaction rates. *Advances in Chemical Physics*, 13:85, 1967.

[37] C. Dykstra, G. Frenking, K. Kim, and G. Scuseria. No Title. In *Theory and Applications of Computational Chemistry. The First Forty Years, Chap. 10*. Elsevier edition, 2005.

[38] Y. Zhang and W. Yang. A challenge for density functionals: Self-interaction error increases for systems with a noninteger number of electrons. *The Journal of Chemical Physics*, 109(7):2604, 1998.

[39] D. R. B. Brittain, C. Y. Lin, A. T. B. Gilbert, E. I. Izgorodina, P. M. W. Gill, and M. L. Coote. The role of exchange in systematic DFT errors for some organic reactions. *Physical Chemistry Chemical Physics : PCCP*, 11(8):1138–42, February 2009.

[40] C. Lee, W. Yang, and R. G. Parr. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical Review B*, 37(2):785, 1988.

[41] A. D. Becke. Density-functional thermochemistry. III. The role of exact exchange. *The Journal of Chemical Physics*, 98(7):5648, 1993.

[42] P. J. Stephens and F. J. Devlin. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *The Journal of Physical Chemistry*, 98(45):11623–11627, 1994.

[43] M. T. Nguyen. Difficulties of density functional theory in investigating addition reactions of the hydrogen atom. *Journal of Physical Chemistry*, 3654(96):18422–18425, 1996.

[44] Y. Zhao, N. González-García, and D. G. Truhlar. Benchmark database of barrier heights for heavy atom transfer, nucleophilic substitution, association, and unimolecular reactions and its use to test theoretical methods. *The journal of physical chemistry. A*, 109(9):2012–8, March 2005.

[45] M. Holzmann, D. Ceperley, C. Pierleoni, and K. Esler. Backflow correlations for the electron gas and metallic hydrogen. *Physical Review E*, 68(4):046707, October 2003.

[46] J. Ma, D. Alfè, A. Michaelides, and E. Wang. The water-benzene interaction: insight from electronic structure theories. *The Journal of chemical physics*, 130(15):154303, April 2009.

[47] F. R. Petruzielo, J. Toulouse, and C. J. Umrigar. Approaching chemical accuracy with quantum Monte Carlo. *The Journal of chemical physics*, 136(12):124116, March 2012.

[48] M. J. Gillan, F. R. Manby, M. D. Towler, and D. Alfe. Assessing the accuracy of quantum Monte Carlo and density functional theory for energetics of small water clusters. *The Journal of chemical physics*, 136(24):244105, June 2012.

[49] D. Alfè, M. Gillan, M. Towler, and R. Needs. Diamond and β-tin structures of Si studied with quantum Monte Carlo calculations. *Physical Review B*, 70(21):1–8, December 2004.

[50] R. Maezono, N. Drummond, A. Ma, and R. Needs. Diamond to β-tin phase transition in Si within diffusion quantum Monte Carlo. *Physical Review B*, 82(18):1–7, November 2010.

[51] S. Sorella, M. Casula, L. Spanu, and A. Dal Corso. Ab initio calculations for the β-tin diamond transition in silicon: Comparing theories with experiments. *Physical Review B*, 83(7):1–12, February 2011.

[52] C. Filippi and C. J. Umrigar. Correlated sampling in quantum Monte Carlo: A route to forces. *Physical Review B*, 61(24):R16291–R16294, June 2000.

[53] S. Chiesa, D. Ceperley, and S. Zhang. Accurate, Efficient, and Simple Forces Computed with Quantum Monte Carlo Methods. *Physical Review Letters*, 94(3):036404, January 2005.

[54] S. Sorella and L. Capriotti. Algorithmic differentiation and the calculation of forces by quantum Monte Carlo. *The Journal of Chemical Physics*, 133(23):234111, December 2010.

[55] A. Badinski, P. D. Haynes, J. R. Trail, and R. J. Needs. Methods for calculating forces within quantum Monte Carlo simulations. *Journal of Physics: Condensed Matter*, 22(7):074202, February 2010.

[56] C. Attaccalite and S. Sorella. Stable Liquid Hydrogen at High Pressure by a Novel Ab Initio Molecular-Dynamics Calculation. *Physical Review Letters*, 100(11):1–4, March 2008.

[57] O. Valsson and C. Filippi. Photoisomerization of Model Retinal Chromophores: Insight from Quantum Monte Carlo and Multiconfigurational Perturbation Theory. *Journal of Chemical Theory and Computation*, 6(4):1275–1292, April 2010.

[58] M. Barborini and L. Guidoni. Reaction pathways by quantum Monte Carlo: insight on the torsion barrier of 1,3-butadiene, and the conrotatory ring opening of cyclobutene. *The Journal of chemical physics*, 137(22):224309, December 2012.

[59] J. Grossman and L. Mitas. High Accuracy Molecular Heats of Formation and Reaction Barriers: Essential Role of Electron Correlation. *Physical Review Letters*, 79(22):4353–4356, December 1997.

[60] C. Filippi, S. Healy, P. Kratzer, E. Pehlke, and M. Scheffler. Quantum Monte Carlo Calculations of H2 Dissociation on Si(001). *Physical Review Letters*, 89(16):166102, September 2002.

[61] G. Mills and H. Jónsson. Quantum and thermal effects in H2 dissociative adsorption: Evaluation of free energy barriers in multidimensional quantum systems. *Physical review letters*, 72(7):1124, 1994.

[62] G. Henkelman, B. P. Uberuaga, and H. Jónsson. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *The Journal of Chemical Physics*, 113(22):9901, 2000.

[63] M. W. Schmidt, K. K. Baldridge, J. a. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery. General atomic and molecular electronic structure system. *Journal of Computational Chemistry*, 14(11):1347–1363, November 1993.

[64] Y. Zhao, B. J. Lynch, and D. G. Truhlar. Multi-coefficient extrapolated density functional theory for thermochemistry and thermochemical kinetics. *Physical Chemistry Chemical Physics : PCCP*, 7:43–52, 2005.

[65] K. A. Peterson and T .H. Dunning Jr. Benchmark Calculations with Correlated Molecular Wave Functions. 11. Energetics of the Elementary Reactions F+ H2, O+ H2, and H'+ HCl. *Journal of Physical Chemistry A*, 5639(97):6280–6292, 1997.

[66] S. Saccani, C. Filippi, and S. Moroni. Minimum Energy Pathways via Quantum Monte Carlo. *J. Chem. Phys.*, 138:084109, 2013.

[67] C. Filippi and C. J. Umrigar. Multiconfiguration wave functions for quantum Monte Carlo calculations of first-row diatomic molecules. *The Journal of Chemical Physics*, 105(1):213, 1996.

[68] M. Burkatzki, C. Filippi, and M. Dolg. Energy-consistent small-core pseudopotentials for 3d-transition metals adapted to quantum Monte Carlo calculations. *The Journal of chemical physics*, 129(16):164115, October 2008.

[69] M. Casula. Beyond the locality approximation in the standard diffusion Monte Carlo method. *Physical Review B*, 74(16):161102, October 2006.

[70] A. J. Williamson, J. C. Grossman, R. Q. Hood, A. Puzder, and G. Galli. Quantum Monte Carlo Calculations of Nanostructure Optical Gaps: Application to Silicon Quantum Dots. *Physical Review Letters*, 89(19):196803, October 2002.

[71] L. K. Wagner and L. Mitas. A quantum Monte Carlo study of electron correlation in transition metal oxygen molecules. *Chemical Physics Letters*, 370(3-4):412–417, March 2003.

[72] E. Batista, J. Heyd, R. Hennig, B. Uberuaga, R. Martin, G. Scuseria, C. Umrigar, and J. Wilkins. Comparison of screened hybrid density functional theory to diffusion Monte Carlo in calculations of total energies of silicon phases and defects. *Physical Review B*, 74(12):121102, September 2006.

[73] E. Sola, J. Brodholt, and D. Alfè. Equation of state of hexagonal closed packed iron under Earths core conditions from quantum Monte Carlo calculations. *Physical Review B*, 79(2):024107, January 2009.

[74] L. K. Wagner and J. C. Grossman. Quantum Monte Carlo Calculations for Minimum Energy Structures. *Physical Review Letters*, 104(21):210201, May 2010.

[75] K. P. Driver, R. E. Cohen, Z. Wu, B. Militzer, P. López Ríos, M D Towler, R J Needs, and J W Wilkins. Quantum Monte Carlo computations of phase stability, equations of state, and elasticity of high-pressure silica. *Proceedings of the National Academy of Sciences of the United States of America*, 107(21):9519–24, May 2010.

[76] R. Q. Hood, P. R. C. Kent, and F. A. Reboredo. Diffusion quantum Monte Carlo study of the equation of state and point defects in aluminum. *Physical Review B*, 85(13):134109, April 2012.

[77] C. J. Umrigar, J. Toulouse, C. Filippi, S. Sorella, and R. G. Hennig. Alleviation of the Fermion-Sign Problem by Optimization of Many-Body Wave Functions. *Physical Review Letters*, 98(11):1–4, March 2007.

[78] M. Nightingale and V. Melik-Alaverdian. Optimization of Ground- and Excited-State Wave Functions and van der Waals Clusters. *Physical Review Letters*, 87(4):043401, July 2001.

[79] S. Sorella. Generalized Lanczos algorithm for variational quantum Monte Carlo. *Physical Review B*, 64(2):024512, June 2001.

[80] E. Neuscamman, C. Umrigar, and G. Chan. Optimizing large parameter sets in variational quantum Monte Carlo. *Physical Review B*, 85(4):045103, January 2012.

[81] C. Adamo and V. Barone. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *The Journal of Chemical Physics*, 110(13):6158, 1999.

[82] Y. Zhao and D. G. Truhlar. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other function. *Theoretical Chemistry Accounts*, 120(1-3):215–241, July 2007.

[83] F. Fracchia, C. Filippi, and C. Amovilli. Size-Extensive Wave Functions for Quantum Monte Carlo: A Linear Scaling Generalized Valence Bond Approach. *Journal of Chemical Theory and Computation*, 8(6):1943–1951, June 2012.

[84] M. Holzmann, B. Bernu, and D. Ceperley. Many-body wavefunctions for normal liquid He3. *Physical Review B*, 74(10):104510, September 2006.

[85] P. López Ríos, A. Ma, N. D. Drummond, M. D. Towler, and R. J. Needs. Inhomogeneous backflow transformations in quantum Monte Carlo calculations. *Physical Review E*, 74(6):066701, December 2006.

[86] D. M. Ceperley. Path integrals in the theory of condensed helium. *Reviews of Modern Physics*, 67(2):279, 1995.

[87] DM Ceperley and EL Pollock. Path-integral computation of the low-temperature properties of liquidˆ{4} He. *Physical review letters*, 56(4):5–8, 1986.

[88] H. F. Trotter. On the product of semi-groups of operators. *Proceedings of the American Mathematical Society*, pages 545–551, 1959.

[89] N. V. Prokof'ev, B. V. Svistunov, and I. S. Tupitsyn. "Worm" algorithm in quantum Monte Carlo simulations. *Physics Letters A*, 9601:253, 1998.

[90] M. Boninsegni, N. V. Prokofev, and B. V. Svistunov. Worm algorithm and diagrammatic Monte Carlo: A new approach to continuous-space path integral Monte Carlo simulations. *Physical Review E*, 74(3):1–16, September 2006.

[91] M. Boninsegni, N. V. Prokofev, and B. V. Svistunov. Worm Algorithm for Continuous-Space Path Integral Monte Carlo Simulations. *Physical Review Letters*, 96(7):070601, February 2006.

[92] N. Metropolis and A. W. Rosenbluth. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21:1087, 1953.

[93] W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57:97, 1970.

[94] E. L. Pollock and D. M. Ceperley. Path-integral computation of superfluid densities. *Physical Review B B*, 36(16):8343–8352, December 1987.

[95] C. J. Umrigar. Accelerated metropolis method. *Physical review letters*, 71(3):408, 1993.

[96] C. J. Umrigar, M. P. Nightingale, and K. J. Runge. A diffusion Monte Carlo algorithm with very small time-step errors. *Journal of Chemical Physics*, 99(4):2865, 1993.